JASA ARTICLE

Gridless three-dimensional compressive beamforming with the Sliding Frank-Wolfe algorithm

Gilles Chardon[1,a)] and Ulysse Boureau[2]

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

[2]*Université Paris-Saclay, CentraleSupélec, Gif-sur-Yvette, 91190, France*

ABSTRACT:

The application of the Sliding Frank-Wolfe algorithm to gridless compressive beamforming is investigated for single and multi-snapshot measurements and the estimation of the three-dimensional (3D) position of the sources and their amplitudes. Sources are recovered by solving an infinite dimensional optimization problem, promoting sparsity of the solutions, and avoiding the basis mismatch issue. The algorithm does not impose constraints on the source model or array geometry. A variant of the algorithm is proposed for greedy identification of the sources. The experimental results and Monte Carlo simulations in 3D settings demonstrate the performances of the method and its numerical efficiency compared to the state of the art. © *2021 Acoustical Society of America*.

https://doi.org/10.1121/10.0006790

(Received 15 June 2021; revised 26 September 2021; accepted 4 October 2021; published online 26 October 2021)

[Editor: Karim G. Sabra] Pages: 3139–3148

I. INTRODUCTION

In the context of source localization,[1] sparsity based estimation methods offer several advantages compared to the classical methods, such as conventional beamforming, mostly in terms of spatial resolution and size of the data.[2] Sparsity based methods can be grossly classified in the optimization based methods, mostly using the ℓ_1 norm,[3–5] or similar mixed norms,[6] and greedy algorithms, the most popular being orthogonal matching pursuit[7] (OMP). These methods accept that the sources are drawn from a finite dictionary of sources and estimate their amplitudes by assuming that most of them are zero. This approach suffers from the basis mismatch problem in which actual sources cannot be exactly represented by a member of the dictionary.[8] It has been shown that this problem cannot be mitigated by refining the grid: for one-dimensional (1D) problems, even in low noise regimes, the least absolute shrinkage and selection operator (LASSO) method yields a number of sources that are doubled compared to the actual source distribution, even if the sources are located on the grid.[9]

Several methods have been proposed to deal with this limitation. The first approach is to approximate a source at an arbitrary point by a source on a finite grid, corrected by an additional term given by a Taylor expansion. The numerical problem is then solved using the sparsity based method[10,11] or sparse Bayesian learning.[12,13]

Grid-free compressive beamforming is possible for far-field sources and uniform linear arrays[14] or uniform rectangular arrays[15] with possibly missing nodes or multiple snapshots.[16] Numerically, a finite dimensional semi-definite program (SDP)[17,18] is used to recover the directions of arrival. The extension to arbitrary array shapes was recently introduced.[19] However, other source models, such as near-field sources, cannot be tackled by this method. Finally, the Newtonized orthogonal matching pursuit (NOMP) algorithm[20,21] is a variant of the OMP, where Newton steps are used to refine the estimations of the sources at each iteration.

In this paper, grid-free source localization is performed by solving the Beurling LASSO problem,[22] an infinite dimensional optimization problem similar to the LASSO problem. The sources are assumed to be located in a region of interest Ω, which is, here, not assumed to be a discrete set. The distribution of sources is modeled by a measure μ defined on Ω. In the particular case of a finite number of monopolar sources, the measure μ is a finite sum of Dirac masses, which are located at the positions of the sources and weighted by their amplitudes. In particular, we consider the Sliding Frank-Wolfe (SFW) algorithm[23] to solve the Beurling LASSO problem and estimate the positions and amplitudes of the sources.

In this algorithm, a source is added at each iteration, and the parameters of all of the sources are then locally and jointly optimized. Under some conditions (in particular, that the solution is a finite sum of Dirac masses and is unique), this algorithm was shown to converge in a finite number of iterations.[23]

In addition to its original formulation, the SFW algorithm is extended here to deal with multi-snapshot data. Moreover, in addition to the standard SFW algorithm, which aims to solve the Beurling LASSO problem, a modification of the SFW algorithm is also used as an OMP algorithm

[a)]Electronic mail: gilles.chardon@centralesupelec.fr, ORCID: 0000-0003-0483-9957.

with local refinements, an alternative to the NOMP. Compared to the SFW algorithm for the Beurling LASSO (BLASSO) problem, this variant takes the number of sources as a parameter, and its number of iterations is limited to the number of sources.

The method is demonstrated on the experimental data with the localization of four sources in a three-dimensional (3D) domain in a case when some sources cannot be identified by beamforming because of their limited resolution. The Monte Carlo simulations are performed to obtain results on the accuracy and computational time of the SFW algorithm compared with previously proposed methods. Although 3D localization, compared to two-dimensional (2D) localization, does not raise fundamental challenges, practical issues arise:[24–26] the distance between the sources and array is an additional parameter to be estimated, the resolution in the radial direction from the array is poorer than in the normal direction and, finally, the size of the domain of interest is increased from a surface to a volume. Therefore, efficient methods, in the sense of computational time, are necessary. The SFW based grid-free methods are shown to be competitive with better or equivalent estimation performances and reduced computational complexity. Moreover, the SFW based methods do not require particular array configurations or source models.

The paper is organized as follows. Section II introduces the source localization model and discusses the state of the art. In Sec. III, the Beurling LASSO and SFW algorithm are recalled, and the multi-snapshot variant is introduced. The numerical and experimental results are given in Secs. IV and V, respectively. Section VI concludes the paper. The code reproducing the numerical and experimental results is available online.[27]

II. MODEL AND STATE OF THE ART

We consider an array of M microphones, located at positions $\mathbf{y}_m \in \mathbf{R}^3$ (m denotes the index of the microphone), measuring acoustical data. The complex amplitudes of the measurements at a given frequency f are obtained at times t_s, where $s = 1,\ldots,S$ is the number of snapshots (in practice, the 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 $\mathbf{p}_s \in \mathbf{C}^M$ at 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. When $S = 1$, we write $\mathbf{p} = \mathbf{p}_1$ and $a_k = a_{k1}$. In the free field conditions, the vector $\mathbf{g}(\mathbf{x})$ is given by its coefficients,

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

where $k = 2\pi f/c$, and c is the wave velocity. However, no particular shape is assumed for $\mathbf{g}(\mathbf{x})$. Our goal is to estimate the positions $\mathbf{x}_k$ and amplitudes a_{ks} from the measured data $\mathbf{p}_s$. The grid-based and grid-free methods are now recalled.

A. Grid-based estimation

In the grid-based methods, the source positions $\mathbf{x}_k$ are assumed to lie on a discrete grid of N points. A finite dimensional dictionary $\mathbf{D} \in \mathbf{C}^{M\times N}$ is then assembled by collecting the Green functions of the grid points, and the vector of the amplitudes of the sources $\mathbf{a} \in \mathbf{C}^N$ is recovered by solving the system

$$\mathbf{p} \approx \mathbf{D}\mathbf{a}. \tag{3}$$

The goal of sparse recovery is to find the solution of Eq. (3) with the least nonzero coefficients, i.e., with the lowest ℓ_0-"norm," defined as the number of nonzero coefficients. This problem is not tractable.

In the OMP, an approximate solution of ℓ_0 minimization is found by selecting the sources iteratively, identifying the source that is the most correlated to the data, and then projecting the data on the orthogonal of the space spanned by the previously identified sources.

In the ℓ_1 based approaches, the ℓ_0-norm is replaced by its convex relaxation, the ℓ_1 norm (or in multiple snapshots settings, a mixed norm[6,28]). A convex minimization problem is then solved, yielding a sparse distribution of sources. The LASSO is a penalized formulation of the problem in which the vector of the estimated amplitudes $\mathbf{a}_\lambda \in \mathbf{C}^N$ is obtained by solving

$$\mathbf{a}_\lambda = \underset{\mathbf{a}\in\mathbf{C}^N}{\operatorname{argmin}} \|\mathbf{D}\mathbf{a} - \mathbf{p}\|_2^2 + \lambda\|\mathbf{a}\|_1, \tag{4}$$

where the ℓ_1 norm is defined as $\|\mathbf{a}\|_1 = \sum_{n=1}^{N} |a_n|$. Alternatively, a constrained version can be used by minimizing an ℓ_1 norm under an ℓ_2 constraint (basis pursuit denoising, BPDN) or vice versa. The application of this method to the acoustical source localization was investigated in Refs. 4 and 25.

The grid-based estimation methods are limited by the basis mismatch problem[8] in which a source does not exactly lie on a grid point. Then, a source is spread out over multiple grid points, hindering the correct estimation of its position and amplitude.

B. Atomic norm based compressing beamforming

In some particular cases, grid-free compressive beamforming can be achieved using finite dimensional positive semi-definite problems.[14,15] This is the case for far-field sources and uniform linear arrays, where $\mathbf{g}(\mathbf{x})$ has a particular shape (that is, complex exponentials). This type of method is limited to specific shapes of microphone arrays and, more importantly, to the far-field case, as it takes advantage of the properties of complex exponentials. This method can be extended to multiple snapshots[16] and arbitrary array shapes by using a Fourier decomposition of the measure modeling the source distribution.[19] However, this method imposes constraints on the model (shape of the array, far-field sources) that will not be satisfied in our numerical and experimental settings as we consider the arbitrary array shapes and non-far-field sources.

C. OMP with local optimization

Starting from the OMP algorithm, the NOMP was proposed for spectral estimation,[20] recently extended to source localization.[21] The idea of the NOMP is to augment the OMP algorithm with local optimization of the positions and amplitudes of the sources using Newton steps. To do so, the sources are added at each iteration on a finite grid and refined using a Newton optimization step, leaving the other positions fixed.

After the introduction of a new source, the positions and amplitudes of all of the sources are refined one at a time, cyclically by local Newton steps, until the decrease in the energy of the residual between the iterations falls below a tolerance τ. A new iteration is then executed with the addition of a new source until a stopping criterion is reached.

The results in the 2D localization showed that estimation of the positions of the sources was improved compared to the OMP, using the same finite grid.[21] However, the method fails with coarse grids when local Newton iterations do not converge toward a local minimum of the objective function.[20]

D. Block-sparsity with Taylor approximations

The block-sparsity, combined with the Taylor approximations, can be used for the grid-free sparse estimation. With $\mathbf{z}$ being a member of the grid, an off-grid source at the position $\mathbf{z}+\Delta$ with the offset $\Delta = (\Delta_x, \Delta_y, \Delta_z)$ of the amplitude a can be approximated as

$$a\mathbf{g}(\mathbf{z}+\Delta) \approx a\mathbf{g}(\mathbf{z}) + a\Delta_x \frac{\partial \mathbf{g}}{\partial x} + a\Delta_y \frac{\partial \mathbf{g}}{\partial y} + a\Delta_z \frac{\partial \mathbf{g}}{\partial z}. \tag{5}$$

For positive real amplitudes a, off-grid sources can be localized by the continuous basis pursuit method,[10] minimization of the ℓ_1 norm of the sources with additional convex constraints on the corrections, ensuring that they remain bounded by $\delta/2$, where δ is the grid step. However, it was shown that for small grid steps δ, this method, as the finite dimensional LASSO, represents a unique source by multiple grid points.[29] For complex amplitudes, the constraints necessary to ensure that the corrections are not larger than $\delta/2$ are non-convex. In this case, a mixed norm can be used.[11] For each grid point, a vector of amplitudes $(\hat{a}_0, \hat{a}_1, \hat{a}_2, \hat{a}_3) \approx (a, a\Delta_x, a\Delta_y, a\Delta_z)$ (most of them begin null) is estimated from which the correction of the position is obtained. However, compared to the positive real amplitudes case, this formulation does not ensure that the corrected point stays near the associated grid point (in fact, the quotients $\hat{a}_{1,2,3}/\hat{a}_0$ are, in general, complex). Finally, the step δ has to be small enough so that the linear approximation around the center of a cell remains precise enough in the cell, implying small steps δ and large computational grids.

III. THE SFW ALGORITHM

We introduce now the Beurling LASSO problem and the SFW algorithm used to solve it. Assuming that the sources are located in a region of interest Ω, the distribution of the sources is modeled by a measure μ, i.e., a function taking as input a subset[30] of Ω, yielding a positive, real, or complex value (respectively, measure, signed measure, and complex measure). The measure μ models the distribution of sources in the domain of interest Ω without the need for a discrete grid.

A particular example of measure is the Dirac mass $\delta_{\mathbf{x}}$, which can be used to model a point source at position $\mathbf{x} \in \Omega$ of unit amplitude. We recall that the integral of a function f with respect to the Dirac mass $\delta_{\mathbf{x}}$ is given by

$$\int_\Omega f d\delta_{\mathbf{x}} = f(\mathbf{x}). \tag{6}$$

In general, a distribution of K sources at time t_s with positions $\mathbf{x}_k$ and complex amplitudes a_{ks} is modeled by the discrete measure

$$\mu_s = \sum_{k=1}^{K} a_{ks} \delta_{\mathbf{x}_k}. \tag{7}$$

Equation (1) can be rewritten as

$$\mathbf{p}_s = \int_\Omega \mathbf{g} d\mu_s + \mathbf{n}_s. \tag{8}$$

The source localization method introduced in this paper is based on the SFW algorithm,[23] which aims at solving the Beurling LASSO problem, which is defined for $S=1$ by

$$\mu^\star = \underset{\mu \in \mathcal{M}}{\operatorname{argmin}} \frac{1}{2} \left\| \int_\Omega \mathbf{g} d\mu - \mathbf{p} \right\|_2^2 + \lambda |\mu|(\Omega), \tag{9}$$

where $\mathcal{M}$ is the set of complex measures defined on Ω. Then, the positions and amplitudes of the sources are given by the decomposition of $\mu^\star$ in Eq. (7).

The regularization term $|\mu|(\Omega)$ is the total variation norm of the complex measure μ, defined by[31]

$$|\mu|(\Omega) = \max_{f \in C_1} \int_\Omega f d\mu, \tag{10}$$

where C_1 is the set of continuous functions on Ω with the absolute value bounded by one. For a discrete measure $\mu = \sum_{k=1}^{K} a_k \delta_{\mathbf{x}_k}$ (with pairwise distinct $\mathbf{x}_k$), the definition reduces to

$$|\mu|(\Omega) = \sum_{k=1}^{K} |a_k|. \tag{11}$$

We observe that here we recover the ℓ_1 norm of the amplitudes of the sources. It can also be considered as the atomic norm of μ using the Dirac masses as building blocks. The Beurling LASSO is similar to the LASSO problem (4) with the important difference that the positions of the estimated sources are not limited to a predefined finite grid.

The SFW algorithm solves Eq. (9) by iteratively adding the Dirac masses to the measure being optimized, alternating with local updates of the positions and amplitudes of the Dirac masses. The steps of the algorithm are given in detail in Algorithm 1. The notation $\mathbf{G}(\mathbf{X})$ denotes the $M\times k$ matrix, collecting the values of $\mathbf{g}$ for the positions in the tuple $\mathbf{X}=(\mathbf{x}_1,\dots,\mathbf{x}_k)$. $(\;)$ denotes the empty tuple, and $(\mathbf{X},\mathbf{x}_\star)$ denotes the insertion of $\mathbf{x}_\star$ into the tuple $\mathbf{X}$. $K_{\max}$ is the maximal number of iterations. An iteration consists of the following steps:

- First, a source is added by solving the global optimization problem (12). To this end, the maximum of $\eta^{[k]}$ [defined in Eq. (13)] is searched on a finite grid and used as the initialization for a local optimization;
- amplitudes are then updated in Eq. (14), which is a LASSO problem;
- finally, the amplitudes and positions are jointly optimized in Eq. (15). This problem is non-convex (indeed, the objective function is left unchanged by the permutation of the positions and amplitudes of two sources, and replacing the positions and amplitudes of these two sources by their mean is unlikely to yield a lower value of the objective function as would be the case for a convex objective function). However, the local optimization is sufficient, e.g., with a quasi-Newton algorithm, initialized at positions $\mathbf{X}^{[k-1/2]}$ and amplitudes $\mathbf{a}^{[k-1/2]}$.

Under some constraints on the solution $\mu^\star$ (in particular, that it is a finite sum of Dirac masses), the algorithm is shown to converge in a finite number of iterations.[23] In our implementation, the MATLAB 2019b function `fmincon` was used to solve Eqs. (12), (14), and (15) using the sequential quadratic programming algorithm.[32]

ALGORITHM 1. The SFW algorithm, solving Eq. (9).

$\mu^{[0]}\leftarrow 0$, $\mathbf{r}^{[0]}\leftarrow\mathbf{p}$, $\mathbf{X}^{[0]}\leftarrow(\;)$

for $k=1,\dots,K_{\max}$ **do**

Identify a new source by solving

$$\mathbf{x}_\star=\operatorname*{argmax}_{x\in\Omega}\eta^{[k]}(\mathbf{x}),\tag{12}$$

where

$$\eta^{[k]}(\mathbf{x})=\frac{1}{\lambda}\left|g(\mathbf{x})^\star\mathbf{r}^{[k-1]}\right|\tag{13}$$

if $\eta(\mathbf{x}_\star)\le 1$ **then**

Stop

else

$\mathbf{X}^{[k-1/2]}=(\mathbf{X}^{[k]},\mathbf{x}_\star)$

Optimize the amplitudes:

$$\mathbf{a}^{[k-1/2]}=\operatorname*{argmin}_{\mathbf{a}\in\mathbf{C}_+^k}\frac{1}{2}\|\mathbf{G}(\mathbf{X}^{[k-1/2]})\mathbf{a}-\mathbf{p}\|_2^2+\lambda\|\mathbf{a}\|_1\tag{14}$$

Optimize the amplitudes and positions:

$$(\mathbf{X}^{[k]},\mathbf{a}^{[k]})=\operatorname*{argmin}_{\mathbf{X}\in\Omega^k,\mathbf{a}\in\mathbf{C}_+^k}\frac{1}{2}\|\mathbf{G}(\mathbf{X})\mathbf{a}-\mathbf{p}\|_2^2+\lambda\|\mathbf{a}\|_1\tag{15}$$

$\mu^{[k]}\leftarrow\sum_{n=1}^{k}a_n^{[k]}\delta_{\mathbf{x}_n^{[k]}}$, $\mathbf{r}^{[k]}\leftarrow\mathbf{p}-\mathbf{G}(\mathbf{X}^{[k]})\mathbf{a}^{[k]}$

end if

end for

A. Multi-snapshots data

In this subsection, we present a modification of the SFW algorithm aimed at processing multi-snapshot data. In the case of several snapshots, the measured data $\mathbf{p}_s$ is assembled in the $M\times S$ matrix $\mathbf{P}$. The amplitudes for each source and each snapshot are optimized in Eqs. (14) and (15). In Eq. (15), the positions of the sources are common among the snapshots. In these equations, the ℓ_1 norm is replaced by a mixed norm.[6,28] With $\mathbf{A}$ being the $k\times S$ matrix containing the complex amplitudes of k sources for S snapshots and each row $\mathbf{a}_{\mathrm{r},j}$ corresponding to the amplitudes of the jth source for the S snapshots, we define the $\ell_{2,1}$ mixed norm as

$$\|\mathbf{A}\|_{2,1}=\sum_{j=1}^{k}\|\mathbf{a}_{\mathrm{r},j}\|_2.\tag{16}$$

In a given row, the ℓ_2 norm is considered, which does not impose sparsity inside the row. Indeed, this would imply temporal sparsity of the source, which is not expected here. Then, the ℓ_1 norm of the ℓ_2 norm is computed as we expect spatial sparsity.

The function η is updated with

$$\eta^{[k]}(\mathbf{x})=\frac{1}{\lambda}\sqrt{\sum_{s=1}^{S}|\mathbf{g}(\mathbf{x})^\star\mathbf{r}_s^{[k-1]}|^2},\tag{17}$$

where $\mathbf{r}_s^{[k-1]}$ is the residual for the sth snapshot, defined by

$$\mathbf{r}_s^{[k]}=\mathbf{p}_s-\mathbf{G}(\mathbf{X}^{[k]})\mathbf{a}_{\mathrm{c},s}^{[k]},\tag{18}$$

where $\mathbf{a}_{\mathrm{c},s}^{[k]}$ is the sth column of $\mathbf{A}^{[k]}$, containing the amplitudes of the sources for the sth snapshot. Details on this choice of η are given in the Appendix.

Equations (14) and (15) are replaced by

$$\mathbf{A}^{[k-1/2]}=\operatorname*{argmin}_{\mathbf{A}\in\mathbf{C}^{kS}}\frac{1}{2}\|\mathbf{G}(\mathbf{X}^{[k-1/2]})\mathbf{A}-\mathbf{P}\|_F^2+\lambda\|\mathbf{A}\|_{2,1}\tag{19}$$

and

$$(\mathbf{X}^{[k]}, \mathbf{A}^{[k]}) = \underset{\mathbf{X}\in\Omega^k, \mathbf{A}\in\mathbf{C}^{kS}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{G}(\mathbf{X})\mathbf{A} - \mathbf{P}\|_F^2 + \lambda\|\mathbf{A}\|_{2,1}, \quad (20)$$

where $||\cdot||_F$ denotes the Frobenius norm of a matrix, defined as the ℓ_2 norm of its coefficients.

B. Setting the regularization parameter λ

Setting the regularization parameter is known to be a difficult problem. To help this tuning, the homotopy approach can be used:[33] results of the algorithm for several parameters λ can be obtained by sequentially running the SFW algorithm for decreasing λ, initializing each run with the output of the previous run. In this case, the algorithm starts by solving Eq. (15) and checking the value of η to avoid the addition of a source if it is not necessary. The search of the parameter λ can be stopped when an appropriate number of sources is found or, according to Morozov's principle,[34] by setting λ such that the error between the measurements and acoustical field generated by the estimated sources $||\mathbf{G}(\mathbf{X})\mathbf{A} - \mathbf{P}||_F^2$ is on the order of the expected noise level.

C. A greedy version of the SFW algorithm

Although the SFW algorithm solves a minimization problem, its structure is similar to the greedy sparse recovery algorithms such as the OMP. The SFW algorithm can be used for greedy identification of acoustical sources by setting $\lambda = 0$, using a normalized dictionary $\mathbf{g}_n(\mathbf{x}) = \mathbf{g}(\mathbf{x})/||\mathbf{g}(\mathbf{x})||_2$, and stopping the algorithm after a fixed number of iterations. This version of the SFW algorithm is more practical when an estimation of the number of sources is known. Moreover, the number of iterations is limited to the iterations necessary to recover the sources, whereas additional spurious sources can slow the algorithm down because of the additional iterations and the higher number of parameters to optimize in Eqs. (14) and (15).

Compared to the NOMP, where positions and amplitudes of the sources are optimized sequentially at each iteration, here, positions and amplitudes of all of the identified sources are locally and jointly optimized. The normalized dictionary $\mathbf{g}_n$ will also be used when solving the BLASSO problem (9).

IV. NUMERICAL RESULTS

Performances of the SFW algorithm for the acoustical source localization, in comparison to the state of the art, are now assessed using simulated data. Both versions of the SFW algorithm are considered: the *penalized SFW*, solving Eq. (9), and the *greedy SFW*, introduced in Sec. III C.

A. A simple 1D example

To allow comparison between the semi-definite programming based method, BPDN, block-sparsity, and penalized SFW algorithm, a first simulation is made in the case of a simple 1D direction of arrival estimation. The results of equivalent source method - iteratively ReWeighted Least-Squares (ESM-IRLS)[24] are also given.

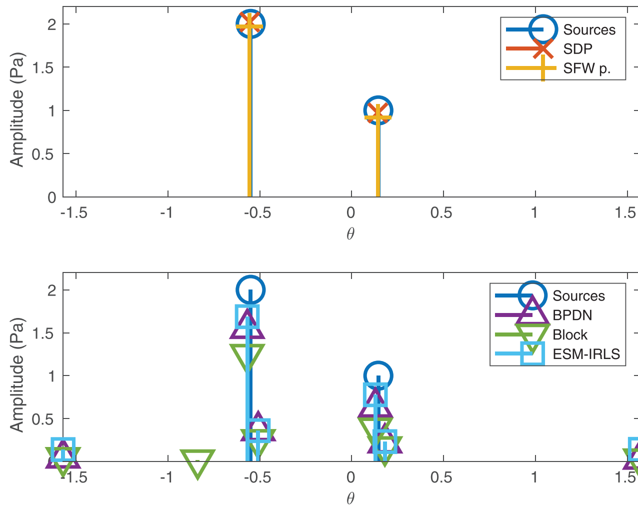

FIG. 1. (Color online) The simulations. The direction of the arrival estimation, (a) SFW and SDP, (b) BPDN and block-sparsity (amplitudes below 10^{-2} are not plotted).

Figure 1 shows the results of the localization of two far-field sources of amplitudes 1 and 2 Pa, arriving from angles $\theta_1 = 0.21$ and $\theta_2 = -0.53$ with a uniform linear array of 20 microphones with a pitch that is half the wavelength, using $S = 1$ snapshot. For the SDP problem, BPDN, and block-sparsity estimation, the tolerance ϵ is set as twice the norm of the noise and for the SFW algorithm, $\lambda = 1$. The ESM-IRLS[24] is used with $p = 0$. The grid used for BPDN, block-sparsity, and ESM-IRLS and to initialize Eq. (12) in the SFW algorithm is a regular sampling of 40 values of $\sin(\theta)$ in the interval $[-\pi/2, \pi/2]$. For the block-sparsity, the derivative of the dictionary is normalized so that the decomposition of a source located between grid points has coefficients in the dictionary and its derivative is of the same order.

Although θ_1 and θ_2 are not on the grid, the SDP and SFW algorithm are capable of estimating the directions of arrival correctly. For the BPDN, block-sparsity, and EMS-IRLS, each source is represented by two spikes. Moreover, several spurious sources appear, caused by the basis mismatch.

B. Performances of the SFW algorithm

Now, performances of the SFW algorithm, in its original penalized version (*SFW p.*) and greedy version (*SFW g.*), OMP, and NOMP are compared in the function of several parameters: step δ of the discretization grid, frequency, SNR, number of snapshots, and resolution using Monte Carlo simulations. The SDP based grid-free method cannot be used here as it cannot deal with the source model. The block-sparsity method was also considered, but long running times (typical running times of 500 s compared to less than a second for the greedy SFW algorithm at $\delta = 0.1$ m) prevented its use in the simulations.

An array of 128 microphones is used with positions of the microphones shown in Fig. 2 in the plane $Z = 0$. The domain Ω is defined by $-1 \leq X \leq 1, -1 \leq Y \leq 1,\ 3 \leq Z \leq 5$ (in m), discretized with a step $\delta = 0.05$ m (except when stated otherwise). Three sources are randomly placed in Ω using a uniform density with amplitudes 0.1, 0.2, and 0.4 Pa and random phases. For the original SFW algorithm, the λ parameter is chosen as the minimal value such that three sources are identified. The estimated amplitudes obtained by least squares fitting of the data to the sources identified by the BLASSO problem, avoiding the bias introduced by the regularization terms, are also given (*SFW p. LS*).

Fifty realizations of the data are used to estimate the mean square error (MSE) of the positions and amplitudes of the sources.

1. Grid step

In the OMP, the estimated positions will necessarily lie on the grid, whereas for the NOMP and SFW algorithm, the grid is used only at the identification step, followed by local optimization, and should have a lesser impact on the localization results.

The MSE for the position and amplitude estimation are given in Fig. 3(a) for the OMP and NOMP with tolerances of $\tau = 10^{-7}$ and $\tau = 10^{-9}$, respectively, and the SFW algorithm with $S = 1$, $F = 4858$ Hz and a SNR of 20 dB. As expected, the performances of the OMP worsen as the step increases. Compared to the OMP, the local optimization of the NOMP with $\tau = 10^{-7}$ improves the localization performances. The NOMP, with $\tau = 10^{-9}$, reaches similar performances to the SFW algorithm for steps up to 0.1 m. However, at $\delta = 0.2$ m, the performances of the NOMP degrade compared to the SFW algorithm. This can be explained by the fact that the condition for convergence of the Newton method toward a local maximum of the likelihood function is not satisfied:[20] at the grid point, where a source is first placed, the objective function is non-convex, and the Newton step does not yield a relevant update of the position. For larger steps, the SFW algorithm also fails. Here, the grid is too coarse compared to the wavelength to allow initialization of the optimization problem (12) near the global optimum.

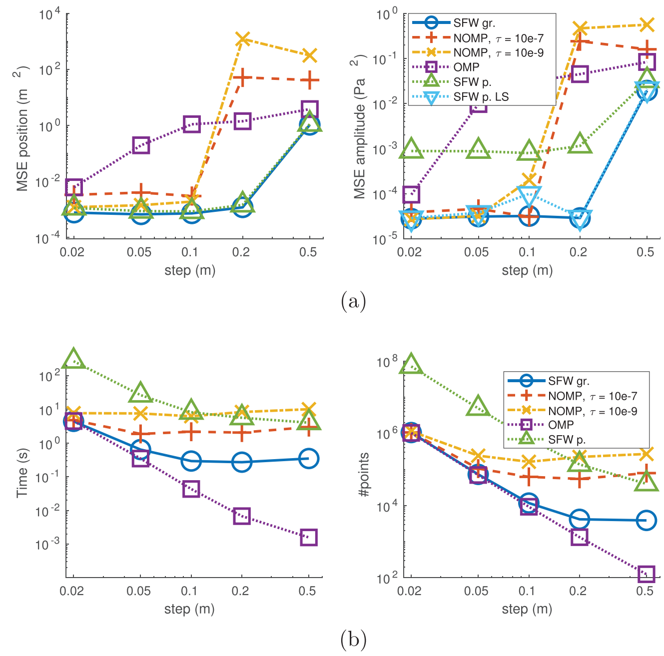

FIG. 3. (Color online) The simulations. The performances of the greedy SFW, penalized SFW, OMP, and NOMP in the function of the grid step. The (a) MSE in position (left) and amplitude (right) and (b) computational time (left) and number of positions visited (right) are shown.

The amplitude estimated by solving the Beurling LASSO problem with the penalized SFW algorithm is biased. Reestimating the amplitudes improves the estimation of the amplitudes.

For the following numerical experiments, the tolerance $\tau = 10^{-9}$ will be used for the NOMP.

2. Frequency

The performances with respect to the frequency are plotted in Fig. 4(a) with $S = 1$, a grid step of 0.05 m, and a SNR of 20 dB. At low frequencies, the methods can return several sources with identical positions when the distance between the sources is not sufficient. In this case, the estimation of the amplitude of the sources is ill-conditioned. To avoid a large error in the amplitude estimations caused by this ill-conditioning, when two sources have estimated positions $\mathbf{x}$ and $\mathbf{x}'$ such that $|\langle \mathbf{g}(\mathbf{x}), \mathbf{g}(\mathbf{x}')\rangle| / (\|\mathbf{g}(\mathbf{x})\| \|\mathbf{g}(\mathbf{x}')\|) > 0.98$, their amplitudes are replaced by their averages and, similarly, when the three sources are estimated at the same position.

Here, it is expected that the performances improve with increasing frequencies. This improvement is visible for all of the methods at lower frequencies (performances of the OMP are, however, limited by the coarse grid used here). Above the frequency 10 kHz for the NOMP and 18 kHz for the SFW algorithm, the performances of the methods

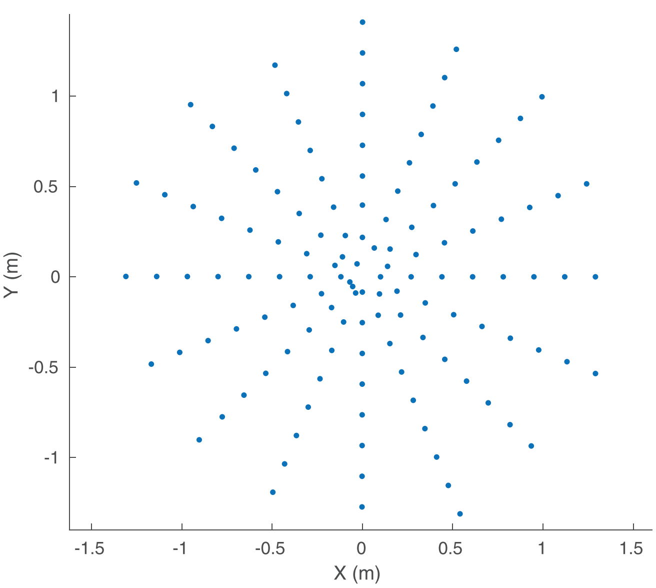

FIG. 2. (Color online) The layout of the microphone array.

https://doi.org/10.1121/10.0006790

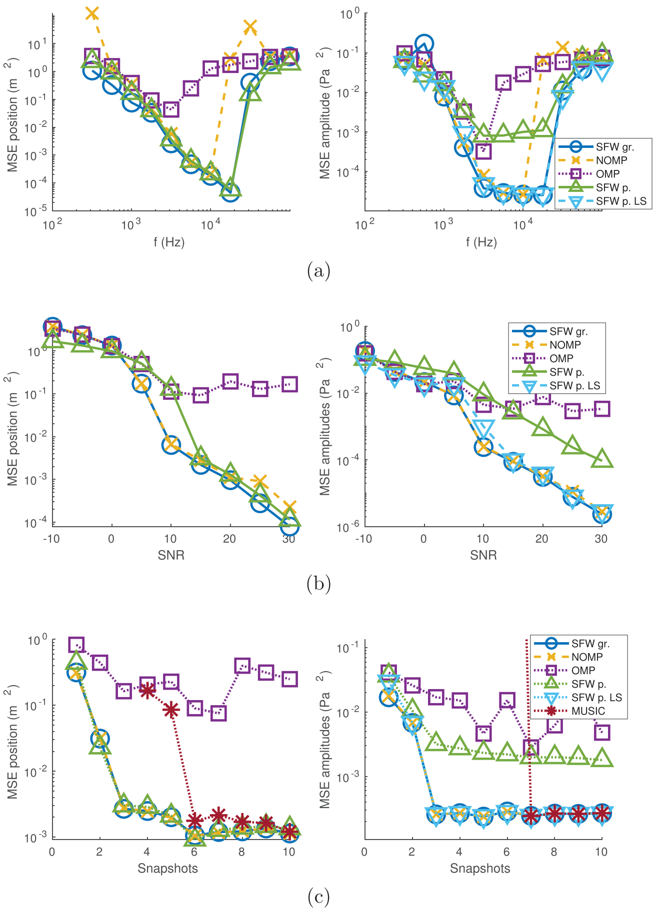

FIG. 4. (Color online) The simulations. The position (left) and amplitude MSE (right) in the function of the (a) frequency F, (b) SNR, and (c) number of snapshots S.

degrade. This failure is explained, as in the case of the varying grid step, by the coarseness of the grid compared to the wavelength.

3. SNR

Here, we analyze the effect of the noise level on the estimation. Figure 4(b) shows the MSE in the position and amplitudes with respect to the SNR with $S = 1$, $F = 4858$ Hz. The performances of the OMP are limited by the grid. The NOMP and greedy SFW algorithm show similar performances except at a high SNR. This discrepancy is caused by stopping the local optimization of the NOMP when the tolerance τ is reached, which is before convergence.

4. Number of snapshots

The performances of the methods are now compared for the increasing number of snapshots from $S = 1$ to $S = 10$ with $F = 4858$ Hz and a SNR of 10 dB in Fig. 4(c). Here, the amplitudes of the sources at each snapshot are drawn from independent random complex Gaussian variables such that their root mean square (RMS) amplitudes are 0.1, 0.2, and 0.4 Pa. All of the methods fail for a unique snapshot. This is caused by the lower SNR and random amplitudes, which, hence, can be close to zero for some sources in some configurations. Using several snapshots ensures that all of the sources have a sufficient amplitude for at least one snapshot. The results of the MUSIC algorithm are also given when the number of snapshots is larger than the number of sources, showing better performances of the SFW algorithm compared to MUltiple SIgnal Classification (MUSIC) for the values of S used here.

5. Resolution

Finally, the resolution of the methods is compared. Two sources of identical amplitudes are used with $S = 1$, $F = 4858$ Hz, and a SNR of 20 dB with varying distances between the sources. Figure 5 shows the position error when the two sources are in a plane parallel to the array (left) and with the two sources having identical X and Y coordinates (right). The greedy SFW algorithm, penalized SWF, and NOMP improve at the same threshold (approximately 0.075 m in the XY plane and 0.44 m in the Z axis). The greedy SFW algorithm exhibits better performances for distances up to twice this threshold.

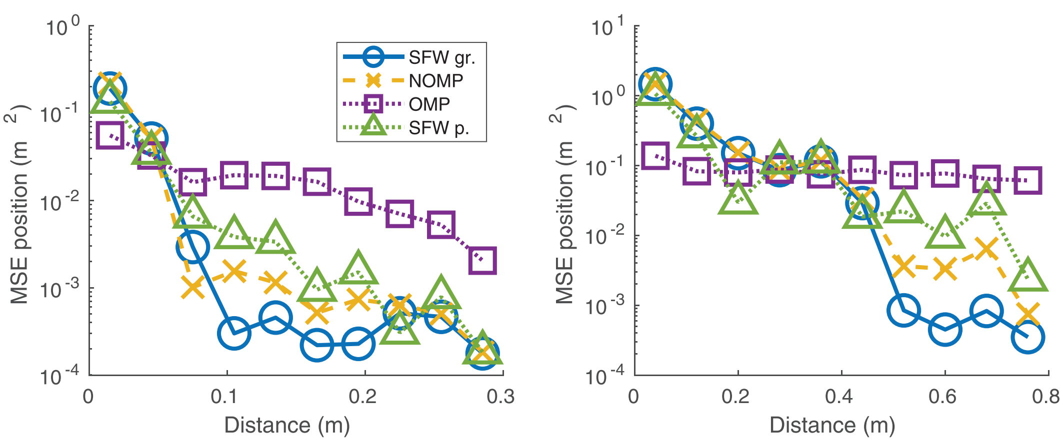

FIG. 5. (Color online) The simulations. The position MSE in the cases of two sources with varying distances are shown with sources in the same plane parallel to the array (left) and in the same axis normal to the array (right).

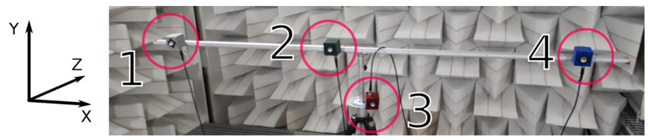

FIG. 6. (Color online) The acoustical sources and coordinate frame.

6. Computational effort

In Fig. 3(b), the computational time [on a personal laptop equipped with an Intel Core i7-7820HQ at 2.90 GHz × 8 central processing unit (CPU) and 16 GB memory; Intel, Santa Clara, CA, USA] of the four methods is plotted, as well as the number of positions $\mathbf{x}$, where $\mathbf{g}(\mathbf{x})$ is computed for a varying grid step. The SFW algorithm yields results in a shorter time than the NOMP at $\tau = 10^{-9}$ by an order of magnitude with fewer visited positions. This discrepancy is, in part, explained by the fact that the SFW algorithm optimizes the positions and amplitudes of all of the sources jointly, whereas the NOMP considers them separately.

In conclusion, these numerical experiments show that the SFW algorithm compares favorably with respect to the NOMP method. Indeed, when both methods succeed in locating the sources, the SFW algorithm yields similar or better performances than the NOMP. Additionally, the SFW algorithm does not necessitate the grids as fine as the NOMP. At a fixed grid size, the NOMP shows higher computational times at a tolerance τ, yielding similar performances to the SFW algorithm. These observations show that the SFW algorithm is better suited than the NOMP for grid-free source localization, in particular, at high frequencies as the necessary grid step scales like the wavelength. The greedy version of the SFW algorithm was found to yield slightly better performances than the SFW algorithm solving the BLASSO problem, even after the reestimation of the amplitudes.

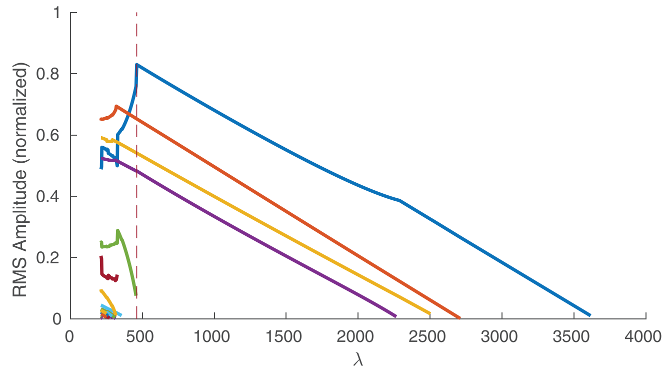

FIG. 7. (Color online) The experimental results. The RMS amplitudes of the sources (identified by a color) found by the penalized SFW with the normalized dictionary $\mathbf{g}_n$ in the function of the regularization parameter λ. The parameter λ used in Fig. 9 is indicated by the dashed line.

V. EXPERIMENTAL RESULTS

In the experimental results, four sources (Visaton-BF32 omnidirectional loudspeakers, Haan, Germany) are used, pictured in Fig. 6, which emit white noise. The acoustical field is measured using an array of 128 Micro Electro Mechanical Systems (MEMS) microphones (InvenSense-INMP441, San Jose, CA, USA) with the positions shown in Fig. 2. The sampling frequency is $F_s = 50$ kHz, and the signals are analyzed by a short time Fourier transform with a Hann window of duration 82 ms (4096 samples) and 50% overlap. More details on the experimental setup are found in Ref. 35.

The domain of interest is the box defined by $-2 \leq X \leq 1, -1 \leq Y \leq 0, 4 \leq Z \leq 5$ (in m), containing the four sources. $S = 10$ snapshots are used, and the results are obtained at frequency $F = 1818$ Hz. The results of conventional beamforming are given in Fig. 9(a). At the chosen frequency, the two central sources cannot be separated by beamforming.

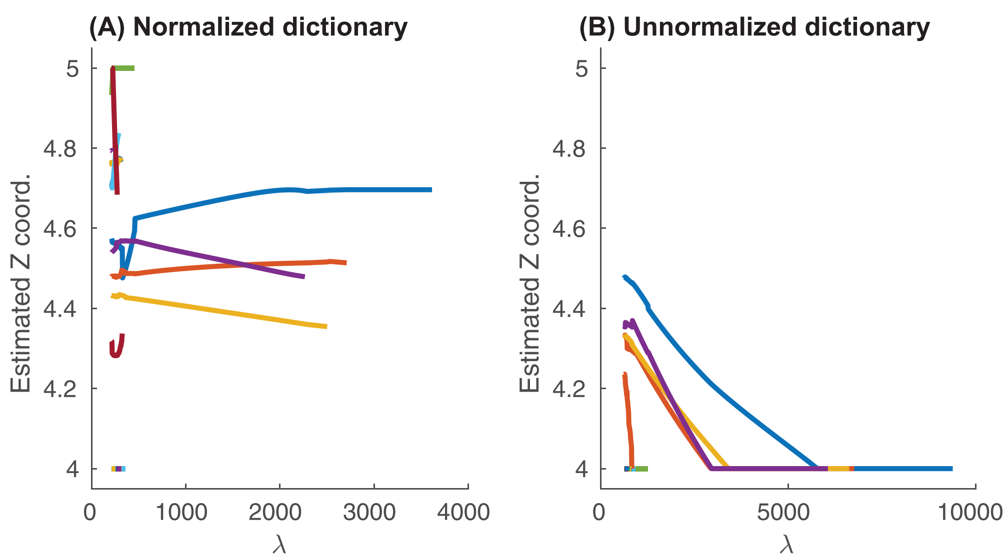

FIG. 8. (Color online) The experimental results. The Z-coordinate of the sources identified by the penalized SFW for the normalized $\mathbf{g}_n$ and un-normalized model $\mathbf{g}$ in the function of the regularization parameter λ.

https://doi.org/10.1121/10.0006790

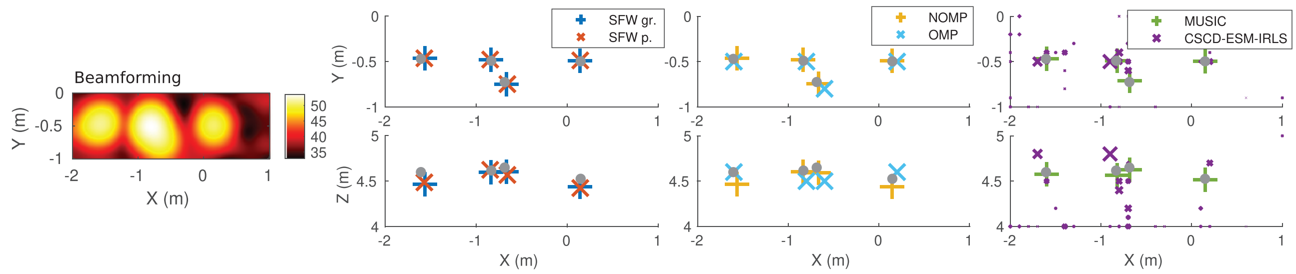

FIG. 9. (Color online) The experimental results. The beamforming in a plane, and greedy and penalized versions of the SFW, OMP, NOMP, MUSIC and CSCD-ESM-IRLS are shown in the front (top) and top (bottom) views. The sizes of the markers are proportional to the source power.

The penalized SFW algorithm is run for several values of λ in the BLASSO problem (9) using the homotopy approach described in Sec. III B with the normalized dictionary $\mathbf{g}_n$. The estimated amplitudes and positions of the sources depend on the regularization parameter λ. Figure 7 shows the evolution of the estimated amplitudes of the sources in the function of λ. The amplitudes are defined as the RMS of the amplitudes for each snapshot, normalized by the amplitude of the most powerful source, and estimated by the beamforming from the measurements where it is the only active source. Each color represents a source identified by the algorithm. For high λ, the amplitudes of the sources are underestimated, whereas for low λ, several spurious sources appear.

Figure 8 highlights the importance of the normalization of the dictionary. Here, the estimated Z coordinates for each source are plotted in the function of the regularization parameter λ, using the un-normalized dictionary $\mathbf{g}$ and the ℓ_2-normalized dictionary $\mathbf{g}_n$. Using $\mathbf{g}$, the estimated positions are biased toward the array for large λ. Indeed, in these cases, it is preferable to estimate a source at a position closer to the array, where the model does not match the data as well as the correct position, but the necessary amplitude is smaller and, therefore, generates a smaller penalization. Although the estimated Z coordinates fluctuate in the function of λ when using $\mathbf{g}_n$, no clear bias is visible.

In Fig. 9, the estimated positions for $\lambda = 648$ (just before the appearance of a fifth source, denoted as the penalized SFW) are plotted. In Fig. 9, the results of the greedy versions of the SFW algorithm, OMP, NOMP (here, the data-dependent tolerance τ is $\tau = 0.01$, which has been observed to yield similar results to the SFW algorithm), MUSIC, and CLEAN-SC Decomposition combined with the ESM-IRLS (Ref. 24) are reported with four iterations.

The estimated powers of the sources are given in Table I for the OMP, NOMP, greedy and penalized SFW, and MUSIC and are compared to the power estimated by beamforming in settings where each source is the unique active source.

TABLE I. Experimental results. The estimated powers of the sources for OMP, NOMP, greedy SFW, penalized SFW (penal. SFW), and penalized SFW with the least squares estimation of the amplitudes (penal. SFW LS).

	Power	OMP	NOMP	SFW gr.	SFW p.	SFW p. LS	MUSIC
Source 1	52.9	51.8	51.6	51.5	50.0	51.6	51.7
Source 2	53.3	53.5	53.4	53.4	52.0	53.4	53.6
Source 3	51.0	48.1	49.6	49.6	47.3	49.5	50.1
Source 4	50.8	50.1	50.1	50.1	48.3	50.1	50.3

VI. CONCLUSION

The application of the SFW algorithm for acoustical source localization is introduced. Modifications of the algorithm were used to take into account the multi-snapshots data and perform greedy identification of the sources. The estimation performances were shown to be better than, or comparable to, the state of the art. Additionally, the method is numerically efficient with smaller computational times than other grid-free methods and is not based on a particular source model or limited to specific array shapes. The results for several values of the regularization parameter λ are obtained, helping the choice of the regularization parameter.

ACKNOWLEDGMENTS

The authors thank François Ollivier (Institut d'Alembert, Sorbonne Université, CNRS) for making the experimental results possible.

APPENDIX A: SELECTION CRITERION η

In Algorithm 1, $\eta^{[k]}$ is defined as

$$\eta^{[k]}(\mathbf{x}) = \frac{1}{\lambda}|\mathbf{g}(\mathbf{x})^\star \mathbf{r}^{[k-1]}|. \tag{A1}$$

At the end of an iteration, where amplitudes and positions of the estimated sources are locally optimal, the objective function can be locally improved only by adding a new source. Defining $h_{\mathbf{x},\psi}(\epsilon) = J(\mu^k + \epsilon e^{i\psi}\delta_{\mathbf{x}}) - J(\mu^k)$, the variation of the objective function J of Eq. (9) when introducing a source at $\mathbf{x}$ with phase ψ and positive amplitude ϵ is

$$h'_{\mathbf{x},\psi}(0) = -\mathrm{Re}\left(e^{-i\psi}\mathbf{g}(\mathbf{x})^\star \mathbf{r}^{[k-1]}\right) + \lambda. \tag{A2}$$

The maximal decrease is obtained when ψ is chosen as the angle $\psi_\star$ of $(\mathbf{g}(\mathbf{x})^\star \mathbf{r}^{[k-1]})^\star$, yielding

$$h'_{\mathbf{x},\psi_\star}(0) = -|\mathbf{g}(\mathbf{x})^\star \mathbf{r}^{[k-1]}| + \lambda. \tag{A3}$$

Maximizing $\eta^{[k]}(\mathbf{x})$ can, therefore, be interpreted as choosing the position, $\mathbf{x}$, where adding a source maximizes the improvement of the objective function. Moreover, when $h'_{\mathbf{x},\psi_\star}(0) \geq 0$ for all possible positions $\mathbf{x}$ [equivalently, $\eta^{[k]}(\mathbf{x}) \leq 1$], the objective function cannot be improved by adding a source, and the algorithm stops.

For the multi-snapshot problem, we define $h_{\mathbf{x},\mathbf{u}}(\epsilon)$ as the variation of the objective function when a source is introduced at $\mathbf{x}$ with amplitudes $u_s\epsilon$ for each snapshot with $||\mathbf{u}||_2 = 1$ and $\epsilon \geq 0$. Then,

$$h'_{\mathbf{x},\mathbf{u}}(0) = -\sum_{s=1}^{S} \mathrm{Re}((u_s\mathbf{g}(\mathbf{x}))^\star \mathbf{r}_s^{[k-1]}) + \lambda. \tag{A4}$$

Defining $\mathbf{v}$ such that $v_s = \mathbf{g}(\mathbf{x})^\star \mathbf{r}_s^{[k-1]}$,

$$h'_{\mathbf{x},\mathbf{u}}(0) = -\mathrm{Re}(\mathbf{u}^\star \mathbf{v}) + \lambda. \tag{A5}$$

The vector $\mathbf{u}$, maximizing the decrease in the objective function for a given position $\mathbf{x}$, is obtained by choosing $\mathbf{u}$ to be colinear with $\mathbf{v}$, that is, $\mathbf{u}_\star = \mathbf{v}/||\mathbf{v}||_2$, and

$$h'_{\mathbf{x},\mathbf{u}_\star}(0) = -||\mathbf{v}||_2 + \lambda. \tag{A6}$$

We then define

$$\eta(\mathbf{x}) = \|\mathbf{v}\|_2/\lambda, \tag{A7}$$

where $\eta(\mathbf{x}) \leq 1$ when the objective function cannot be decreased by adding a source.

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," IEEE Signal Process. Mag. **13**(4), 67–94 (1996).
[2] P. Gerstoft, C. F. Mecklenbräuker, W. Seong, and M. Bianco, "Introduction to compressive sensing in acoustics," J. Acoust. Soc. Am. **143**(6), 3731–3736 (2018).
[3] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," SIAM Rev. **43**(1), 129–159 (2001).
[4] P. Simard and J. Antoni, "Acoustic source identification: Experimenting the ℓ_1 minimization approach," Appl. Acoust. **74**(7), 974–986 (2013).
[5] A. Xenaki, P. Gerstoft, and K. Mosegaard, "Compressive beamforming," J. Acoust. Soc. Am. **136**(1), 260–271 (2014).
[6] D. Malioutov, M. Cetin, and A. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," IEEE Trans. Signal Process. **53**(8), 3010–3022 (2005).
[7] Y. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers* (1993), Vol. 1, pp. 40–44.
[8] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," IEEE Trans. Signal Process. **59**(5), 2182–2195 (2011).
[9] V. Duval and G. Peyré, "Sparse regularization on thin grids I: The Lasso," Inverse Problems **33**(5), 055008 (2017).
[10] C. Ekanadham, D. Tranchina, and E. P. Simoncelli, "Recovery of sparse translation-invariant signals with continuous basis pursuit," IEEE Trans. Signal Process. **59**(10), 4735–4744 (2011).
[11] Y. Park, W. Seong, and P. Gerstoft, "Block-sparse two-dimensional off-grid beamforming with arbitrary planar array geometry," J. Acoust. Soc. Am. **147**(4), 2184–2191 (2020).
[12] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," IEEE Trans. Signal Process. **61**(1), 38–43 (2013).
[13] A. Das, "Deterministic and Bayesian sparse signal processing algorithms for coherent multipath directions-of-arrival (DOAs) estimation," IEEE J. Ocean. Eng. **44**, 1150–1164 (2019).
[14] A. Xenaki and P. Gerstoft, "Grid-free compressive beamforming," J. Acoust. Soc. Am. **137**(4), 1923–1935 (2015).
[15] Y. Yang, Z. Chu, Z. Xu, and G. Ping, "Two-dimensional grid-free compressive beamforming," J. Acoust. Soc. Am. **142**(2), 618–629 (2017).
[16] Y. Park, Y. Choo, and W. Seong, "Multiple snapshot grid free compressive beamforming," J. Acoust. Soc. Am. **143**(6), 3849–3859 (2018).
[17] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," Commun. Pure Appl. Math. **67**(6), 906–956 (2014).
[18] C. Fernandez-Granda, "Super-resolution of point sources via convex programming," Inf. Inference **5**(3), 251–303 (2016).
[19] Z. Chu, Y. Liu, Y. Yang, and Y. Yang, "A preliminary study on two-dimensional grid-free compressive beamforming for arbitrary planar array geometries," J. Acoust. Soc. Am. **149**(6), 3751–3757 (2021).
[20] B. Mamandipoor, D. Ramasamy, and U. Madhow, "Newtonized orthogonal matching pursuit: Frequency estimation over the continuum," IEEE Trans. Signal Process. **64**(19), 5066–5081 (2016).
[21] Y. Yang, Z. Chu, Y. Yang, and S. Yin, "Two-dimensional Newtonized orthogonal matching pursuit compressive beamforming," J. Acoust. Soc. Am. **148**(3), 1337–1348 (2020).
[22] Y. de Castro and F. Gamboa, "Exact reconstruction using Beurling minimal extrapolation," J. Math. Anal. Appl. **395**(1), 336–354 (2012).
[23] 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 (2020).
[24] G. Battista, P. Chiariotti, M. Martarelli, and P. Castellini, "Inverse methods in aeroacoustic three-dimensional volumetric noise source localization and quantification," J. Sound Vib. **473**, 115208 (2020).
[25] F. Ning, J. Wei, L. Qiu, H. Shi, and X. Li, "Three-dimensional acoustic imaging with planar microphone arrays and compressive sensing," J. Sound Vib. **380**, 112–128 (2016).
[26] T. Padois and A. Berry, "Two and three-dimensional sound source localization with beamforming and several deconvolution techniques," Acta Acust. Acust. **103**, 392–400 (2017).
[27] G. Chardon and U. Boureau, "gilleschardon/SFWCB," at https://zenodo.org/record/5528801 (Last viewed 26/9/2021).
[28] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," IEEE Trans. Signal Process. **54**(12), 4634–4643 (2006).
[29] V. Duval and G. Peyré, "Sparse spikes super-resolution on thin grids II: The continuous basis pursuit," Inverse Problems **33**(9), 095008 (2017).
[30] Technically, it is a member of a σ-algebra of Ω. For the sake of clarity, such technicalities will not be considered here and are discussed at length in appropriate textbooks (see Ref. 31).
[31] W. Rudin, *Real and Complex Analysis*, 3rd ed. (McGraw-Hill, New York, 1987).
[32] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. (Springer, New York, 2006), Chap. 18.
[33] J.-B. Courbot and B. Colicchio, "A fast homotopy algorithm for gridless sparse recovery," Inverse Problems **37**(2), 025002 (2021).
[34] O. Scherzer, "The use of Morozov's discrepancy principle for Tikhonov regularization for solving nonlinear ill-posed problems," Computing **51**(1), 45–60 (1993).
[35] G. Chardon, F. Ollivier, and J. Picheral, "Localization of sparse and coherent sources by orthogonal least squares," J. Acoust. Soc. Am. **146**(6), 4873–4882 (2019).
[36] Technically, a member of a σ-algebra of Ω. For the sake of clarity, such technicalities will not be considered here and are discussed at length in appropriate textbooks (see Ref. 31).