

# A blind dereverberation method for narrowband source localization

Gilles Chardon, Thibault Nowakowski, Julien de Rosny, Laurent Daudet

▶ **To cite this version:**

Gilles Chardon, Thibault Nowakowski, Julien de Rosny, Laurent Daudet. A blind dereverberation method for narrowband source localization. IEEE Journal of Selected Topics in Signal Processing, 2015, pp.11. 10.1109/JSTSP.2015.2422673 . hal-01148974

**HAL Id: hal-01148974**

**https://inria.hal.science/hal-01148974**

Submitted on 6 May 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JSTSP.2015.2422673



# A blind dereverberation method for narrowband source localization

Gilles Chardon, *Member, IEEE*, Thibault Nowakowski, Julien de Rosny, Laurent Daudet, *Senior Member, IEEE*

***Abstract*—Narrowband source localization gets extremely challenging in strong reverberation. When the room is perfectly known, some dictionary-based methods have recently been proposed, allowing source localization with few measurements. In this paper, we first show that, for these methods, the choice of frequencies is important as they fail to localize sources that emit at a frequency near the modal frequencies of the room. A more difficult case, but also important in practice, is when the room geometry and boundary conditions are unknown. In this setup, we introduce a new model for the acoustic soundfield, based on the Vekua theory, that allows a separation of the field into its reverberant and direct source contributions, at the cost of more measurements. This can be used for the design of a dereverberation pre-processing step, suitable for a large variety of standard source localization techniques. We discuss the spatial sampling strategies for the sound field, in order to successfully recover acoustic sources, and the influence of parameters such as number of measurements and model order. This is validated in numerical and experimental tests, that show that this method significantly improves localization in strong reverberant conditions.**

***Index Terms*—Source localization, microphone array, reverberation.**

## I. Introduction

Acoustic or electromagnetic source localization is an inverse problem for which numerous methods have been developed, based on various models and algorithms. A common assumption is that propagation occurs in an open, or anechoic, environment. In this case, the Green function of the medium is known and the direct problem can be solved in a straightforward way. In this paper, we tackle the problem of source localization in an unknown room (e.g. its shape and/or the boundary conditions are unknown). Although not enough data is at hand to solve the direct problem (i.e. the computation of the acoustical field emitted by the sources), we show that using an adequate measurement scheme and preprocessing of the data, it is possible to use standard source localization methods (e.g. beamforming, MUSIC, sparse recovery, etc.). Of course, this possibility comes at a cost, namely a larger microphone array that has to enclose the domain of interest.

There has been an increased activity in the last few years, where localization in this framework was improved by explicitly taking into account the specifics of the environment, e.g. the shape of the room and the reflective properties of its walls, assuming these are known [1], [2]. The goal of this paper is to compare the performance of this approach with known environment, to a more generic model of reverberation introduced here, based on a model of the wavefield that does not require the knowledge of the propagation environment. We first recall some existing methods for different cases: known or unknown environment, and time or frequency domain.

In the case of known environments and in the time domain, by using the fact that the wave equation is invariant to reversing the time parameter, so-called time reversal techniques [3], [4] allow a robust localization of one source. After recording the sound radiated by a source on an array of transducers, these recordings are played backwards by the microphones. The resulting soundfield (that can be either produced experimentally or simulated) focuses back to the location of the original source. However, its resolution is limited by the standard wave diffraction limit, and this method does not easily take into account prior knowledge on the source.

In [5], a sparse modelization of the soundfield in space and time is introduced, and used to localize the sources. The impulse response between the potential source and the microphones is assumed to be known and computed via the image source method. This method is therefore limited to cases where the room is explicitly known and the impulse responses can be computed with sufficient accuracy.

Another method in the known environment/time domain case is the use of cosparsity [6]. The pressure field created by a low number $N$ of sources in a domain $\Omega_0$ is solution to the wave equation with Neumann boundary conditions (in the ideal case of rigid walls)

$$\begin{cases} \frac{\partial^2 p}{\partial t^2} - c^2 \Delta p = \sum_{j=1}^{N} s_j \delta_{x_j} \\ \frac{\partial p}{\partial n} = 0 \text{ on } \partial\Omega_0 \end{cases} \tag{1}$$

where $c$ is the wave velocity, and $s_j$ the sound emitted by the $j$-th source, located at $x_j$. The normal derivative is denoted by $\partial/\partial n$. After discretization (using e.g. finite differences), this can be interpreted as a cosparse model [7] for the pressure, allowing the recovery of the positions of the sources and the signals they emit.

In the frequency domain case, Dokmanic and Vetterli proposed a method allowing the localization of sources in a known reverberant environment using measurements in multiple frequency bands [1]. They replaced the free-field impulse response of the sources by the impulse response computed using the finite element method, and used this as the array manifold, or dictionary. They used a multichannel Orthogonal Matching Pursuit (OMP) to consider the different frequency

Copyright (c) 2014 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org

GC is with Supélec, Gif-sur-Yvette, France. TN, JdR and LD are with Institut Langevin, ESPCI ParisTech, Paris Diderot University, CNRS UMR 7587, Paris, France.

Copyright (c) 2015 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JSTSP.2015.2422673



bands. As will be shown in this paper, this method needs more and more frequency bands as the number of sources increases, and the choice of these frequencies is critical. A variant of this method in the time domain was proposed in [2].

For the localization in an unknown environment and using narrowband measurements, we introduce a model for sound fields based on the Vekua theory [8], [9], and the associated signal processing. This processing removes the contributions of the reverberation, as well as of sources outside of the domain of interest. Once the sound field is processed, classical source localization methods can be applied (beamforming, MUSIC, sparse methods, etc.).

In section II, generalities about source localization are recalled. Section III considers the use of a dictionary, computed from a physical model, when room geometrical and physical properties are known. The main result of the paper, a preprocessing step for source localization in unknown reverberant environment, is introduced in IV and tested in section V, with simulated and real measurements. Concluding remarks are given in section VI.

## II. Source localization

In the case of free-field propagation in the harmonic regime, the acoustical field created at the position $x$ by a source located at $y$ at frequency $\omega$ is given by the free-field Green function

$$G_0(y,x) = \frac{i}{4} H_0(k\|x-y\|) \tag{2}$$

in 2D, where $H_0$ is the outgoing Hankel function of order 0, and $k$ the wavenumber, with $k = \omega/c$, or

$$G_0(y,x) = \frac{\exp(ik\|x-y\|)}{4\pi\|x-y\|} \tag{3}$$

in 3D. The field radiated by $J$ sources located at positions $y_j$ with complex amplitudes $a_j$ is thus

$$p(x) = \sum_{j=1}^{J} a_j G_0(y_j, x). \tag{4}$$

The goal of source localization is to estimate, from a finite number of measurements of $p$, the positions $y_j$ of the sources.

Using $\mathbf{g}_0(y) = (G_0(y,x_i))_i$, the vector of the pressure radiated by sources located at positions $y_j$ with complex amplitudes $a_j$ measured at the positions $x_i$ is given by

$$\mathbf{p} = \sum_j a_j \mathbf{g}_0(y_j). \tag{5}$$

If we further assume that the points $y_j$ are located on a predefined grid of $L$ points $z_l$, we can build a dictionary $\mathbf{G}_0$ of possible sources, with $\mathbf{g}_0(z_l)$ as the $l$-th column. Using this dictionary, we can write

$$\mathbf{p} = \mathbf{G}_0 \mathbf{a} \tag{6}$$

where $\mathbf{a}$ is a $L$-dimensional vector.

The size of $\mathbf{a}$ is likely to be larger than the number of measurements, making the recovery of $\mathbf{a}$ impossible without prior information. However, as only $J$ coefficients of $\mathbf{a}$ corresponding to the actual sources are nonzero, $\mathbf{a}$ is sparse. Practically, this means that only a few sources are present in the domain of interest.

Numerous methods exist to estimate the positions of the sources. Here, we briefly introduce some of the most commonly-used methods : beamforming, (sparsity-based) Orthogonal Matching Pursuit, and MUSIC.

Beamforming estimates the position of a source by correlating measurements with a dictionary of signals emitted by candidate sources. With $\mathbf{p}$ the measurements on a microphone array and $\mathbf{u}(x)$ the field radiated by sources located on a set of possible locations $x$ and measured at the positions of the microphones, the output of the beamformer is

$$\mathbf{b}(x) = \frac{|\mathbf{u}(x)^\star \mathbf{p}|}{\|\mathbf{u}(x)\|} \tag{7}$$

Large values of $\mathbf{b}$ correspond to possible sources.

Orthogonal Matching Pursuit (OMP) can be seen as a refinement of this technique. Sources are identified iteratively. The first source is estimated at the maximum $x_1$ of $b$. Its contribution is then removed from the measurements by projecting on the orthogonal complement of $\mathbf{u}(x_1)$. The second source is then estimated at the maximum $x_2$ of the correlations with the projected measurements, and the measurements projected on the orthogonal complement the space spanned by $\mathbf{u}(x_1)$ and $\mathbf{u}(x_2)$, iterating until enough sources have been identified.

MUSIC is based on the construction of a signal-subspace, i.e. the subset spanned by a set of measurements (*snapshots*) taken at different times (assuming of course that the sources did not move). Sources are located by considering the projection of $\mathbf{u}(x)$ on the orthogonal complement of this subspace. If a source is present at a given position $x_1$, $\mathbf{u}(x_1)$ is included in the signal subspace, and its projection on the orthogonal complement is null. This method has the advantage of having a high resolution, but needs multiple snapshots and uncorrelated sources.

These methods use, implicitly or explicitly, the sparsity of $\mathbf{a}$ to estimate the positions of the sources. A critical point of this framework, common to all source localization methods, is the computation of the dictionary and its numerical properties. While the construction of the dictionary is straightforward in the free-field case (it is given by the free-field Green function (2) or (3)), its computation in the case of a room is more involved, and the case of the unknown room requires a more complex model. These two cases are considered in the next two sections.

## III. Sparsity-based localization in a known room

In the case of source localization in a reverberant room $\Omega$, the Green function no longer has the simple form of the free-field case given in (3), but is dependent on the shape of the room and on the boundary conditions. For the sake of simplicity, we will consider ideal rigid walls, i.e. Neumann boundary conditions on the boundary of the domain of interest $\partial\Omega$. To treat this case, Dokmanic and Vetterli suggest replacing the free-field dictionary by an ad-hoc dictionary computed a priori. The sound field emitted by a source at location $y$ is solution to the Helmholtz equation with a right hand side and

Copyright (c) 2015 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JSTSP.2015.2422673



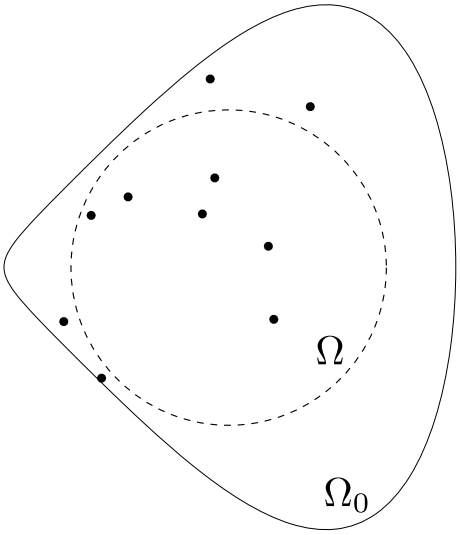



Fig. 1. Shape $\Omega_0$ of the room used for the numerical experiments. Black dots indicates the measurement points for the beamforing results of Fig. 2. $\Omega$ is the domain of interest for the unknown room case.

Neumann boundary conditions:

$$\begin{cases} \Delta p + k^2 p = \delta_y \\ \frac{\partial p}{\partial n} = 0 \text{ on } \partial\Omega \end{cases} \tag{8}$$

By solving this equation for a set of positions of the source (e.g. a regular sampling of the domain where sources are expected), we can build a dictionary for the localization of the sources in this particular room. This dictionary can be used with different methods, e.g. beamforming or, as in [1], iterative algorithms based on sparsity.

We will show that, in this framework:

- it is actually necessary to use measurements at multiple frequencies to locate more than one source,
- the choice of the frequencies is critical.

To illustrate this, we simulate the propagation in the room pictured in Fig. 1, described by the parametric equations

$$\begin{cases} x = \cos t \\ y = \sin t + \frac{1}{3}\sin 2t \end{cases} \qquad t \in [0, 2\pi) \tag{9}$$

The dimensions of this room are approximately 2×2.3 (arbitrary units). This generic shape provides a complex non-separable wavefield (as opposed to rectangular rooms), and allows a straightforward computation of the wave propagation (in particular, no singular corners are present).

We use the FreeFem++ software (based on the finite element method) to simulate the data, and the Vekua-Bergman method to compute the dictionary [10]. Two different methods for the simulation and the construction of the dictionary are used to avoid the "inverse crime" of using the same numerical model for the simulation and the inversion. Although any numerical method can be used to compute the dictionary, the Vekua-Bergman method is numerically efficient and simple to implement.

This method is based on the decomposition of the solution $p$ of (8) as

$$p = p_s + p_0 \tag{10}$$

where $p_s$ is the free-field Green function $G_0(x, y)$ (see (2) or (3)) and $p_0$ the solution to the homogeneous Helmholtz equation

$$\begin{cases} \Delta p_0 + k^2 p_0 = 0 \\ \frac{\partial p_0}{\partial n} = -\frac{\partial p_s}{\partial n} \text{ on } \partial\Omega \end{cases} \tag{11}$$

Such homogeneous solutions to the Helmholtz equation can be approximated by finite sums of plane waves or Fourier-Bessel functions:

$$p_0 \approx \sum_{l=-L}^{L} \alpha_l e^{i\vec{k}_l \cdot \vec{x}} \text{ or } p_0 \approx \sum_{l=-L}^{L} \beta_l J_l(kr) e^{in\theta} \tag{12}$$

where the wave-vectors $\vec{k}_l$ are chosen uniformly on the circle of radius $k$ in the wavenumber domain, and $J_l$ is the $l$-th Bessel function. This was first shown by Vekua for the Fourier-Bessel case [8]. Based on the Vekua theory, convergence of these approximations is studied in [11]. For both cases, the best approximation error between $p_0$ and its approximation of order $L$ is bounded by

$$\|p_0 - p_L\|_{H^m} \leq L^{-(n-m-\epsilon)} \|p_L\|_{H^n} \tag{13}$$

where $\epsilon$ is any strictly positive number, and norms are taken in the Sobolev spaces $H^n$ and $H^m$. When $p$ is smooth, the convergence in $L_2$-norm is exponential. In practice, an order $L \approx kD$, where $D$ is the diameter of the domain, is sufficient. The coefficients of this approximation of $p_0$ are estimated by fitting the boundary conditions of (11) in the least-squares sense on a sampling of the boundary. In practice, the vector $\boldsymbol{\alpha}$ of coefficients is obtained by $\boldsymbol{\alpha} = -\mathbf{W}^\dagger \mathbf{p_{bc}}$, where $\mathbf{W}^\dagger$ is the Moore-Penrose pseudo-inverse of the matrix $\mathbf{W}$ containing the values of normal derivatives of the plane waves or Fourier-Bessel functions on the discretization of the boundary, and $\mathbf{p_{bc}}$ the vector containing the normal derivatives of $p_s$ on the discretization.

The normalized output of the classical beamformer is plotted in Fig. 2 for two sources indicated by the white circles, and 10 sensors (their positions are indicated on Fig. 1). White indicates an output of zero, black the estimated presence of a source. On the left column, only one frequency is used: from the top to bottom, a modal frequency of the room ($k = 9.16$), a random frequency ($k = 9.29$, and a frequency exactly between two modal frequencies ($k = 9.21$). Using a modal frequency makes the localization of the source impossible. Indeed, the field $p$ radiated by a source at position $y$ can be expanded in the modal basis of the room:

$$p(x) = \sum_{s \in \mathbf{N}} \frac{u_s(x) u_s(y)}{k^2 - k_s^2} \tag{14}$$

where $k_s$ are the eigenfrequencies of the room and $u_s$ the corresponding eigenmodes.

It should be noted that, near some eigenfrequency $k_{s_0}$, the mode $u_{s_0}$ (if the eigenfrequency is non-degenerate) dominates the field. Considering this mode only would make source localization impossible, as moving the source has the same effect as changing the amplitude of the mode. In practice, there are always some small contributions of modes other than $u_{s_0}$, either due to the fact that we are not exactly at some modal frequency, or due to losses in the resonant room.

Copyright (c) 2015 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JSTSP.2015.2422673



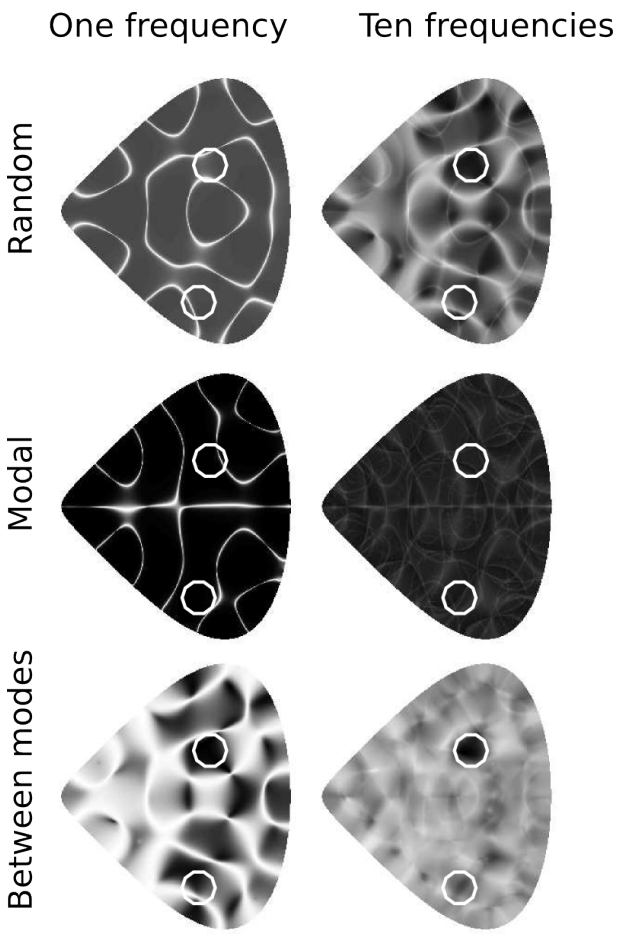



Fig. 2. Known room - Localization of two sources (indicated by white circles) with 10 measurements. Result of the beamformer for random frequencies (top), modal frequencies (middle), means of modal frequencies (bottom). Left column: one frequency, right column: 10 frequencies.

In this case, the possibility of localizing sources depends on the ability of measuring the contribution of these other modes within the total field. This becomes more and more difficult as we approach the modal frequency (or, equivalently, as losses get smaller).

Using frequencies far from modal frequencies, in particular the mean of two successive modal frequencies, yields more informative results, but localization of the two sources is still difficult. On the right column, ten frequencies are used (between $k = 9.16$ and $k = 11$ for the three choices of frequencies), and the output of the ten beamformers are summed. Here also, localization using modal frequencies is impossible. However, with non-modal frequencies, the position of the two sources can be estimated from the output of the beamformers. Using means of successive modal frequencies yields a better contrast.

These observations are supported by the following simulation results in the context of sparse recovery as is done in [1]. In order to evaluate the localization problem in a known room, comprehensive simulations are run with varying numbers of sources, microphones, frequencies, and three choices of frequencies:

- random draw of frequencies within a given interval,
- modal frequencies of the room,
- means of two successive modal frequencies.

Localization results for 2 sources are shown in Fig. 3. The experiment is repeated 20 times, for a number of measurements from 1 to 25, and a number of frequencies from 1 to 30. Frequencies are chosen in the interval $k = 4$ to $k = 10$. A source is considered localized if an estimated source is at a distance less than $\epsilon = 0.15$. The respective performances of the three choices of frequencies are clearly different, and coherent with the results of Fig. 2. These results highlight clear differences in the three possible strategies for choosing the frequencies: while the random choice has mediocre performance and the use of the modal frequencies does not yield exploitable results whatever the numbers of measurements and frequencies, using frequencies between these modal frequencies makes the localization possible, although a large number of frequencies is needed to achieve robust localization. Even in this case, using too few frequencies prevents the localization of the sources.

This result, in the sparse recovery context, can be explained as follows. As shown in Eq. (14), the effect of the position of the source is to change the amplitude of the contributions of the modes. Near an eigenfrequency, the contribution of the mode associated with the eigenfrequency dominates the acoustical field, and the acoustical field is always almost proportional to this mode. As the columns of the dictionary are normalized, they are very similar one to each other. A consequence on the dictionary is that, if the measurements are not all on nodal lines, its coherence (i.e. the maximal scalar product between normalized columns. Note that this definition is unrelated to the concept of wave coherence in physics) tends to 1 as the frequency approaches an eigenfrequency. This is detrimental to the localization, as the coherence of the dictionary has to be low to ensure reconstruction by a sparse recovery algorithm such as OMP [12]. In our case, it means that it is in practice impossible to locate more than one source with only one frequency, even with a larger number of microphones.

In summary, these results show that the propagation in a reverberant room, even when explicitly taken into account by the localization method, has a great impact on the performances of source localization methods. In particular, when using the dictionary computed from the physical model of the room, multiple frequencies are needed, and these frequencies have to be chosen carefully.

## IV. Unknown room

We now turn to the case of an unknown room, for which we do not know the shape and/or the boundary conditions. This makes the direct problem, and thus the computation of the dictionary, impossible to solve. We propose an alternative modelization of the soundfield radiated by sources in a reverberant environment, which allows localization from narrowband measurements in unknown rooms using a simple preprocessing step combined with standard source localization methods.

Copyright (c) 2015 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JSTSP.2015.2422673



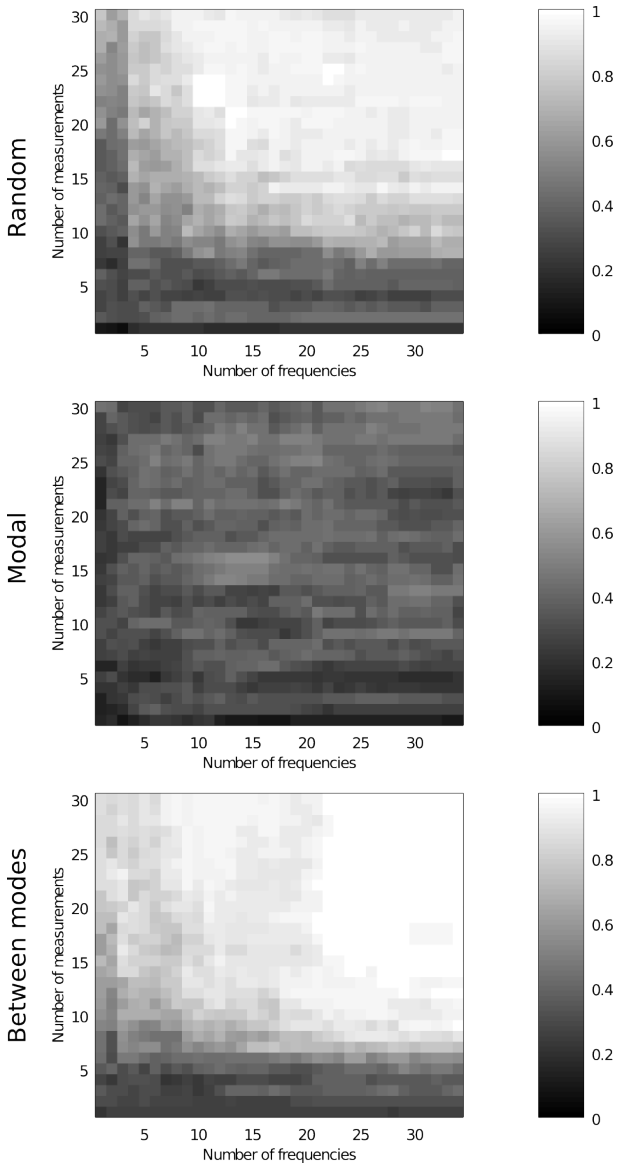



Fig. 3. Known room - Probability of successful localization for the three different choices of frequencies, with 2 sources, for 20 trials as a function of the number of measurements and the number of frequencies. a) random draw of the frequencies, b) modal frequencies of the room, c) means of two successive modal frequencies. Black = no source localized, white = all sources localized. Tolerance $\epsilon = 0.15$.

### A. Dereverberation

We assume that the soundfield is measured in a domain of interest $\Omega$, contained in the room $\Omega_0$ (see Fig. 1). The pressure $p$ radiated by the sources is solution to the Helmholtz equation with a right hand side, and boundary conditions which are unknown to us:

$$\Delta p + k^2 p = s \text{ in } \Omega. \tag{15}$$

The right hand side $s$ contains the sources inside $\Omega$, and has to be identified. Despite the unknown boundary conditions, the above-presented models (10) and (12) still hold for the approximation of the sound field.

When applying a standard source localization algorithm with a free-field model to the measurements of the pressure, localization of the source may become difficult because of the mismatch between the physical model and the dictionary. As it is impossible to construct a dictionary adapted to the physical model without knowledge of the boundary conditions, our method will be based on the removal of the contribution of the reverberation to the measurements.

The pressure $p$ can be written as

$$p = \sum_j a_j g_0(y_j) + p_0, \tag{16}$$

where $g_0$ is a fundamental solution to the Helmholtz equation (e.g. the free-field Green function (2) or (3)), and $p_0$ a solution to the Helmholtz equation without a right hand side, such that the boundary conditions are satisfied.

To make the localization possible, we project the measurements and the dictionary of sources on the orthogonal complement of the solutions to the Helmholtz equation without sources (in which $p_0$ is contained). This projection eliminates the unknown contribution from the boundary conditions (i.e. the reverberation), entirely included in the component $\mathbf{p}_0$.

Note however that sources inside the domain $\Omega_0$ but outside of $\Omega$ will not appear in the right-hand side of this equation when it is limited to the domain $\Omega$. In the above-presented model (10), their contribution is limited to the homogeneous component $p_0$, and will also be eliminated with this component. This can be viewed both as a feature of the method, which is able to neglect unwanted sources outside of its domain of interest, but also as a disadvantage, as it makes localization of sources possible only within this domain.

In practice, the field is measured at a finite number of positions in the domain $\Omega$. We simply project the measurements on the orthogonal complement of the space spanned by plane waves or Fourier-Bessel functions. The orthogonal projector is denoted by $\Pi$, and obtained as $\Pi = \mathbf{W}\mathbf{W}^+$, where $\mathbf{W}$ is the matrix containing the Fourier-Bessel functions or plane waves sampled at the measurements points, and $\mathbf{W}^+$ is its Moore-Penrose pseudo-inverse. Before projection, we have

$$\mathbf{p} = \sum_j a_j \mathbf{g}_0(y_j) + \mathbf{p}_0, \tag{17}$$

where the term $\mathbf{p}_0$ makes the localization impossible. After projection, we obtain

$$\Pi\mathbf{p} = \sum_j a_j \Pi\mathbf{g}_0(y_j) + \Pi\mathbf{p}_0. \tag{18}$$

As the homogeneous term $p_0$ can be approximated by a combination of plane waves or Fourier-Bessel function, $\Pi\mathbf{p}_0$ is small when enough Fourier-Bessel functions are used (as stated above, the norm of $\Pi\mathbf{p}_0$ decays exponentially with the order of approximation, and is low enough for practical purposes when $L \geq kD$). As the field created by a source located inside $\Omega$ is not a solution to the homogeneous Helmholtz in this domain, its contribution is modified, but not eliminated, by the projection $\Pi$.

In this discussion, the free-field Green function was used. Note however that two different choices of Green functions actually yield the same projected dictionary, as the difference between two fundamental solutions centered at the same point

Copyright (c) 2015 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JSTSP.2015.2422673



is a solution to the Helmholtz equation without right hand side and is eliminated by the projection.

Standard localization techniques (beamforming, MUSIC, sparse recovery, etc.) can then be used, with $\Pi\mathbf{G}_0$ as the dictionary and $\Pi\mathbf{p}$ as the measurements. The equation (6) becomes:

$$\Pi\mathbf{p} \approx \Pi\mathbf{G}_0\mathbf{a}. \tag{19}$$

Using the projection $\Pi$ imposes some constraints on the measurements. As already pointed out, the measurements have to enclose the domain of interest $\Omega$. Otherwise, the domain of interest can be restricted to a smaller domain $\Omega'$ enclosing the measurements, where the field created by a source located in $\Omega\backslash\Omega'$ is a solution to the homogeneous equation in $\Omega'$ and is eliminated from the measurements. Moreover, in eq. (18), the vectors have $N$ coordinates, but as they are in the orthogonal complement of the space spanned by the $2L+1$ Fourier-Bessel functions used to capture the homogeneous field, they belong to a $N-(2L+1)$-dimensional subspace of $C^N$. Compared to a free-field problem, $2L+1$ additional measurements are necessary.

## V. Numerical and experimental results

### A. Numerical results

We now test different source localization methods:

- beamforming
- MUSIC
- Orthogonal Matching Pursuit, for which extensive results are given.

The objective of these numerical experiment is not to compare the merit of these three localization methods, but to show that the proposed pre-processing step is compatible with them. We use the same domain as above, but restrict the domain of interest to a disk $\Omega$ of diameter $R = 1.4$, in which the sources are randomly drawn.

The outputs of beamforming and MUSIC, as well as the positions estimated by OMP, are given in Fig. 4 for two sources, with 100 measurements, $k = 20$, 71 Fourier-Bessel functions, a signal to noise ratio of 20dB and two snapshots (i.e. the acoustical field is measured at two different times). Without pre-processing (left column), the localization of the sources is impossible, while after removal of the homogeneous component of the measurements, an accurate localization of the sources is possible (right column), even when no information about the room is available.

Figure 5 shows the probability of localization of 4 sources for OMP and MUSIC as a function of the number of measurements and of Fourier-Bessel functions used to capture the reverberation. A source is considered localized if an estimated source lies less that $\epsilon = 0.1$ from it. The numerical experiment is repeated 20 times with different source positions and 4 snapshots. These results shows that if enough Fourier-Bessel functions are used to capture the reverberation and sufficient measurements are available (more than the number of Fourier-Bessel functions), localization of sources in an unknown reverberant environment is possible. The measurements are here chosen from the border of $\Omega$ as well as inside.

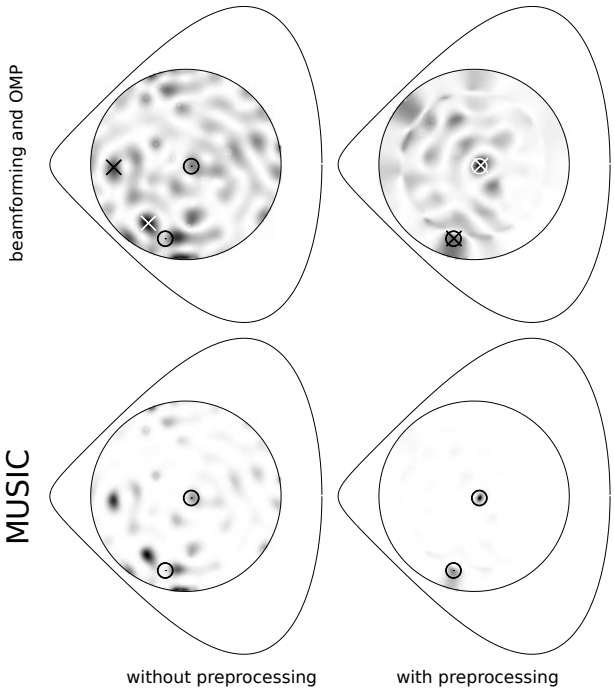



Fig. 4. Unknown room - Output of the beamforming (top) and MUSIC (bottom) methods. The real positions of the sources are indicated by circles, the positions estimated by OMP are indicated by crosses. On the left, a free-field propagation model is assumed, on the right, the homogeneous part is removed from the measurements and the dictionary.

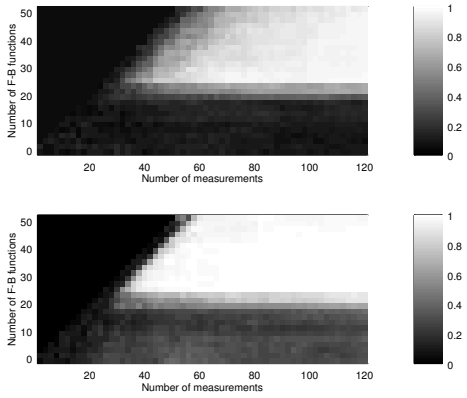



Fig. 5. Unknown room - Probability of successful localization for OMP (top) and MUSIC (bottom), with 4 sources and $k = 15$, for 20 trials as a function of the number of measurements and the number of Fourier-Bessel functions. Black = no source localized, white = all sources localized. Tolerance $\epsilon = 0.1$.

The effects of the choice of the distribution of the samples, the number of measurements and Fourier-Bessel functions, as well as the frequencies are now investigated. As we can see in Fig. 5, OMP and MUSIC have similar behaviors in function of the number of measurements and number of Fourier-Bessel functions. We will thus limit our numerical experiments to the OMP algorithm. Like above, we give, for different sets of parameters, the empirical probability of localizing a source (we considered that a source is localized if an estimated source lies at a distance smaller that $\epsilon = 0.1$).

Copyright (c) 2015 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JSTSP.2015.2422673



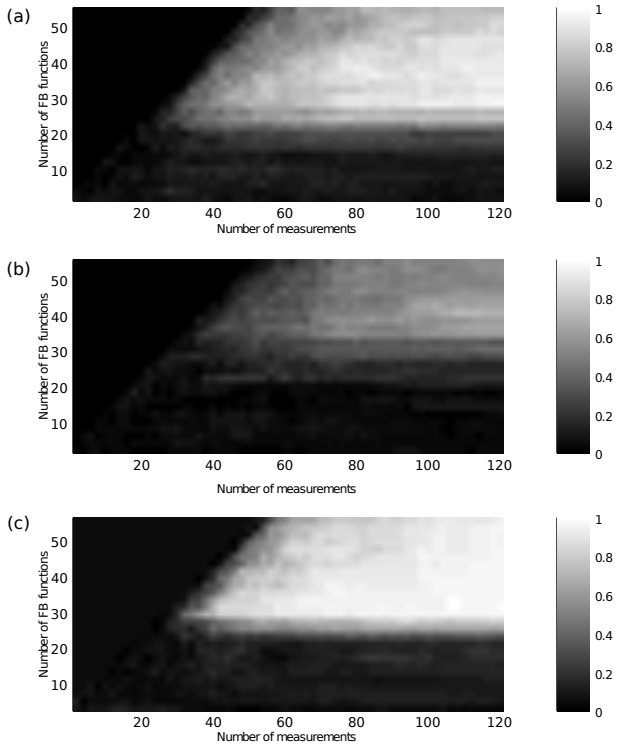



Fig. 6. Unknown room - Probability of successful localization for the three different sampling densities, with 2 sources and $k = 15$, for 20 trials as a function of the number of measurements and the number of Fourier-Bessel functions. a) uniform density in the domain, b) uniform density on its border, c) mixture between these densities. Black = no source localized, white = all sources localized. Tolerance $\epsilon = 0.1$.

*a) Distribution of the samples:* We first test three different sampling strategies, for which the sampling points are drawn using three different probability densities:

- uniform density in the domain,
- uniform density on its border,
- 50% in the domain, 50% on its border.

Results for these 3 densities are given in Fig. 6, for the case of two sources with $k = 15$.

Sampling on the border only fails, as it is actually impossible to distinguish the field created by a source from a homogeneous field using only measurements on the border. Indeed, if one or more sources radiate a pressure $p_s$ on the border, the solution to

$$\begin{cases} \Delta p_0 + k^2 p_0 = 0 \\ p_0 = p_s \text{ on } \partial\Omega_0 \end{cases} \tag{20}$$

is a homogeneous field with the same value on the border of $\Omega$.

Mixed sampling has slightly better performances than interior sampling, in particular for high numbers of Fourier-Bessel functions. This is likely due to the fact the higher order Fourier-Bessel functions are better identified with mixed sampling, as is shown in [13].

*b) Number of measurements and Fourier-Bessel functions:* We here explain the particular shape of the domain of parameters for which the method works.

In Fig. 7, we compare the performance for varying number of sources and measurements with $k = 10$ and $N_{fb} = 21$,

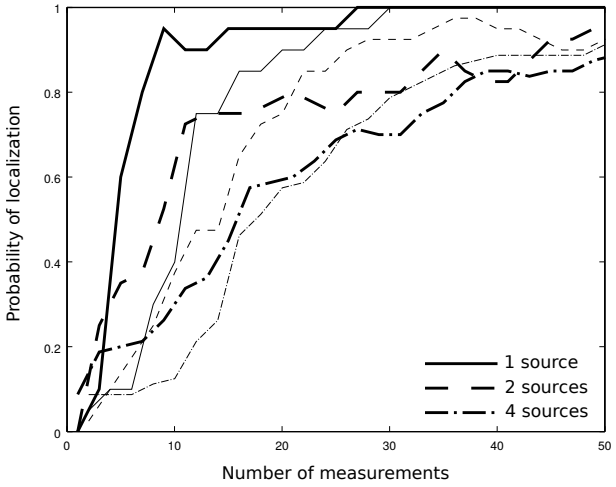



Fig. 7. Unknown room - Probability of successful localization with varying number of sources and number of measurements. $k = 10$, 21 Fourier-Bessel functions. Results are given for a the reverberant room (thick lines) as well as free-field (thin lines). For the reverberant case, the number of measurements indicated is the actual number of measurements minus the number of Fourier-Bessel functions.

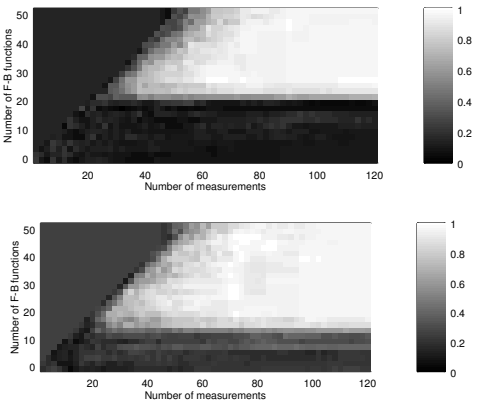



Fig. 8. Unknown room - Probability of successful localization with 2 sources for a modal frequency ($k = 9.98$) and a frequency between two modes ($k = 10.08$)

in the reverberant and free-field cases. For the reverberant case, the number of measurements indicated is the actual number of measurements minus the number of Fourier-Bessel functions. We see that when using this corrected number of measurements for the reverberant case, the behavior of the source localization method is similar for the reverberant and free-field cases.

*c) Wavenumber:* In the known room case, the particular structure of the dictionary makes the localization possible only when using multiple frequencies that are not modal frequencies of the room. The proposed method for the unknown room case is less sensitive to the frequency of the measurements. Results of the proposed method are given in Fig. 8 for two different frequencies, an eigenfrequency and a frequency between two modes. These frequencies are chosen close enough so that the main difference in behavior is due to their modal or not character, and not to their respective

Copyright (c) 2015 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JSTSP.2015.2422673



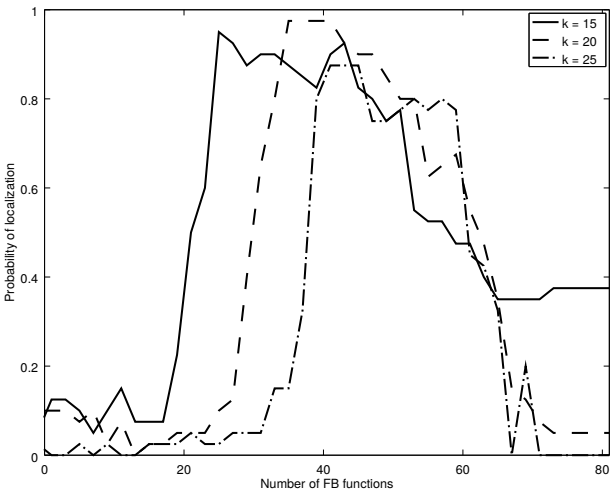



Fig. 9. Unknown room - Probability of successful localization with 2 sources for $k = 15, 20, 25$, 70 measurements and varying number of Fourier-Bessel functions.

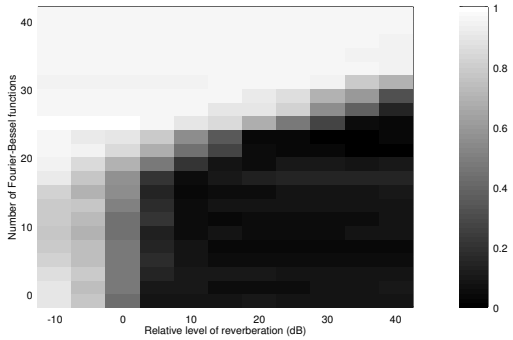



Fig. 10. Unknown room - Probability of successful localization with 50 measurements at $k = 15$, with varying ratio between the energy of the direct field and the energy of the reverberated field, and varying number of Fourier-Bessel functions.

magnitude. The main difference between the two cases is that a larger number of Fourier-Bessel functions has to be used to capture the homogeneous part in the case of a modal frequency. This is expected as the homogeneous part has more energy in this case, and high-order components that were not large enough to perturb the localization have to be eliminated. This obviously makes the minimal number of measurements higher, but not unreasonably so. While some differences can be seen in Fig. 8, the overall performance is similar, if slightly better for the non-modal case. While the choice of the frequency is not as critical as in the case of the known room, it is still preferable to use frequencies between modes to locate sources in this case.

We test, for different wavenumbers ($k = 15, 20, 25$) and fixed number of measurements ($N_m = 70$) and sources ($N_s = 2$), the performance of the method as a function of the number of Fourier-Bessel functions. As seen in Fig. 9 the minimal number of Fourier-Bessel functions required to localize sources depends on the wavenumber. This minimal number is approximately $kD$ where $D$ is the diameter of $\Omega_0$. Using too many Fourier-Bessel functions makes the localization impossible, as no information is left after projection on the orthogonal complement of the Fourier-Bessel functions.

*d) Effect of the relative level of the reverberated field:* As seen in Fig. 8 and 9, the probability of localization of the sources depends on the frequency. More precisely, the number of Fourier-Bessel functions needed for localization is higher for higher frequencies, but also for frequencies, where the reverberated field has more energy.

Here we study the effect of the energy of the reverberated field. We plot in Fig. 10 the probability of localization of two sources as a function of the number of Fourier-Bessel functions and the relative level of the reverberated field in dB. Different levels are obtained by amplifying or attenuating the reverberation in the numerical simulation. In the simulations, the free-field component $p_s$ and the reverberation $p_0$ can be separated. We then simulate different levels of reverberation by changing the amplitude of $p_0$: $p' = p_s + \alpha p_0$ where $\alpha$ is chosen to yield the desired reverberation to direct path ratio. While this method does not correspond to an actual physical setting, it has the advantage of considering only a change in the level of the reverberation, and not a change in its spatial shape.

The bottom row of the figure, corresponding to the use of no FB function, shows the result of standard OMP, without taking reverberation explicitly into account. While localization of the sources is possible when the level of the reverberation is lower than the direct field with OMP and a dictionary of free-field sources, the reverberated field has to be removed from the measurements when its relative level starts to be roughly at the same level, or higher, than the level of the direct field. The distance between a source and a microphone for which the ratio is 1 is called the critical distance [14], and depends on the physical properties of the room. When the microphone array is circular or spherical, the average distance between the source and the microphones is approximately the radius of the array. As a bottom line, separating out the reverberant field, as proposed in this study, is necessary when the size of the array gets larger than the critical distance.

It is also visible that the more the reverberated field has energy, the more Fourier-Bessel functions have to be used. This is expected, as the energy captured by these functions is an increasing function of their number. Using enough Fourier-Bessel functions allows the localization of the sources even if the homogeneous field has an energy several orders of magnitude higher than the direct field of the sources. This can happen when sources outside the domain of interest have to be removed from the measurements.

*e) Computed dictionary vs. projected dictionary:* In cases where the room is known, it is of course also possible to use this blind method. On Fig. 11, we compare the results of OMP with a computed dictionary (using the boundary conditions), and OMP with projected dictionary and measurements (with 31 Fourier-Bessel functions), using 5 to 80 measurements, at $k = 10.51$ (a modal frequency), and $k = 10.34$ (between modes), for 4 sources. Localization is repeated 20

Copyright (c) 2015 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JSTSP.2015.2422673



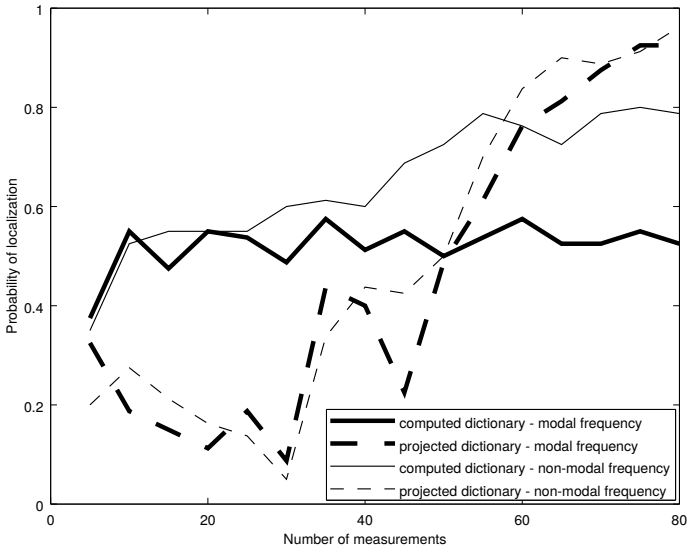



Fig. 11. Result of the localization of 4 sources with computed dictionary or projected dictionary at $k = 10.51$ (a modal frequency), and $k = 10.34$ (between modes), for 5 to 80 measurements.

times with different sources and measurements positions. The probability of localizing a source at a distance $\epsilon = 0.05$ is plotted. Remarkably, robust localization is achieved by the blind method, while using the actual dictionary is limited to the localization of half of the sources. This can be explained by the fact that projecting the dictionary on the orthogonal complement of the Fourier-Bessel functions eliminates (or strongly attenuates) the modes with high amplitude that make localization difficult. While the computed dictionary has a very high coherence (close to 1), the projected dictionary is less coherent and makes the localization possible. We however do not claim that prior information on the room is useless, but that such construction of a source dictionary is unable to benefit from this information.

### B. Experimental results

We here give experimental results for the localization of a source in a strongly reverberating room. The room has dimensions 5.7m $\times$ 4m $\times$ 3m, and a TR60 of 4s. Using the Sabine formula, the critical distance is estimated to be 25cm. As the antenna is a cube of width 1m, the typical distance between a source and a microphone is larger than the critical distance. The correlations between the complete field, and the field after removal of the homogeneous part are given in Fig. 12 for one source.

For the removal of the homogeneous field, we used the version of Eq. (11) obtained by using spherical harmonics and spherical Bessel functions instead of circular harmonics and Bessel functions, or using plane waves with wave-vectors on the sphere of radius $k$. Theoretical convergence results in this 3D setting are similar to the 2D setting. Spherical harmonics are used here for their slightly better numerical properties.

42 microphones are used, with a sampling frequency of 96 kHz. The positions of the microphones are estimated using the method described in [15]. The correlations are computed

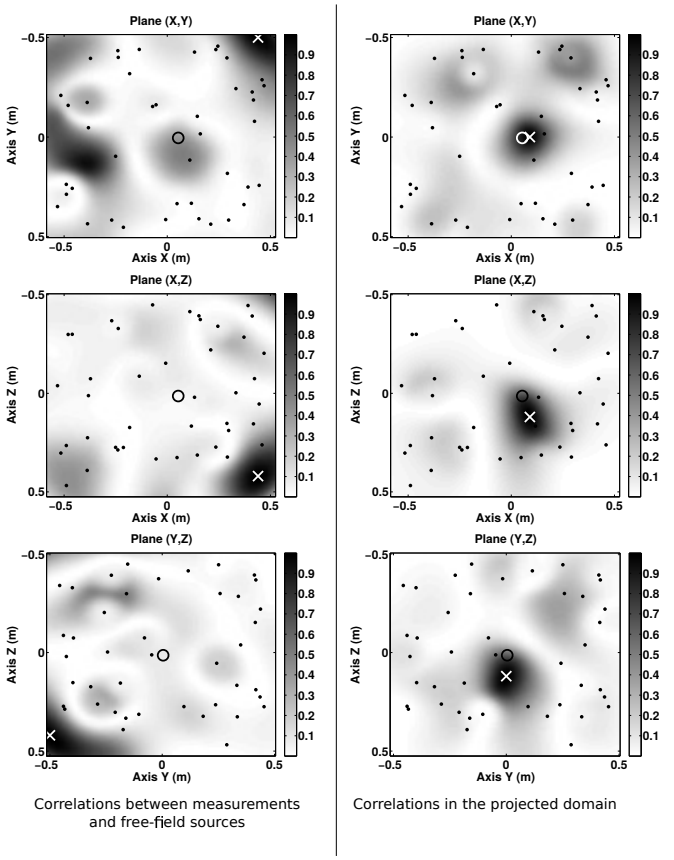



Fig. 12. Unknown room, experimental results. Correlations of the source dictionary with the total field (left) and the field after removal of the homogeneous part (spherical harmonics of order 5, 36 functions) (right) on the three planes in which the source is located. The source is indicated by a circle, its estimation by a cross. The projection of the microphones on the three planes are indicated by dots. Measurements obtained in a 5.7m $\times$ 4m $\times$ 3m room at 628 Hz.

at 628 Hz (i.e. a wavelength of 0.54cm), and 36 functions are used to approximate the homogeneous field (in 3D, products of spherical Bessel functions and spherical harmonics of order up to 5). The energy of the estimated homogeneous component is 12dB higher than the energy of the estimated source component. The left column of Fig. 12 shows, in the 3 directions, the correlation of the source dictionary and the measured field (i.e the modulus of the scalar product between the measurements and the possible sources sampled at the measurements points, after preprocessing or not, and normalized). Without further processing, estimating the source position fails. The right column of Fig. 12 displays the same correlations, after removal of the homogeneous components. Its global maximum is now very close to the actual source position. Experimental validation using more sources, and in other rooms, will be the topic of further studies.

## VI. Conclusion

Our experiments confirm that narrowband localization of sources in known or unknown reverberant rooms is possible using adequate models. However, the two cases have quite different requirements in terms of measurements. While the

Copyright (c) 2015 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JSTSP.2015.2422673



known room case can deal with a small number of measurements, the numerical properties of the dictionaries require the use of multiple and carefully chosen frequencies. More precisely, we show that the measurements should be done between the modal frequencies of the room.

Our main result is the design of a preprocessing step for source localization methods, able to remove the reverberation from acoustical measurements without any prior on the room or on the signals. This preprocessing step is compatible with various localization methods, and makes the localization of sources in a unknown reverberant environment possible using only one frequency. However, more measurements than in the free-field or known room cases are needed, as some of them are used to separate the direct response from the reverberation. Another difference is that this scheme can only localize sources surrounded by measurements. It is shown that this method is beneficial to the localization when the size of the array is larger that the critical distance of the room, and is demonstrated on preliminary experiments.

## ACKNOWLEDGMENTS

GC is supported by the Austrian Science Fund (FWF) START-project FLAME ("Frames and Linear Operators for Acoustical Modeling and Parameter Estimation"; Y 551-N13). LD is on a joint affiliation with Institut Universitaire de France. LD, JdR and TN are supported by LABEX WIFI under references ANR-10-LABX-24 and ANR-10-IDEX-0001-02 PSL*. GC performed part of this work at Institut Langevin and the Acoustic Research Institute, Austrian Academy of Sciences, Vienna, Austria. TN is supported by the Direction Générale de l'Armement.

The authors wish to thank Olivier Lombard for his help in the acquisition of the experimental data.

## REFERENCES

[1] I. Dokmanic and M. Vetterli, "Room helps: acoustics localisation with finite elements," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing, Kyoto, Japan*, 2012.

[2] J. LeRoux, P. T. Boufounos, K. Kang, and J. R. Hershey, "Source localization in reverberant environments using sparse optimization," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, (Vancouver, Canada), May 26-31 2013.

[3] M. Fink, "Time reversal of ultrasonic fields. i. basic principles," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 39, no. 5, pp. 555–566, 1992.

[4] R. K. Ing, N. Quieffin, S. Catheline, and M. Fink, "In solid localization of finger impacts using acoustic time-reversal process," *Applied Physics Letters*, vol. 87, no. 20, p. 204104, 2005.

[5] D. Model and M. Zibulevsky, "Signal reconstruction in sensor arrays using sparse representations," *Signal Processing*, vol. 86, no. 3, pp. 624 – 638, 2006.

[6] S. Nam and R. Gribonval, "Physics-driven structured cosparse modeling for source localization," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing, Kyoto, Japan*, 2012.

[7] S. Nam, M. Davies, M. Elad, and R. Gribonval, "The cosparse analysis model and algorithms," *Applied and Computational Harmonic Analysis*, vol. 34, no. 1, pp. 30 – 56, 2013.

[8] I. N. Vekua, *New methods for solving elliptic equations*. North-Holland, 1967.

[9] P. Henrici, "A survey of I. N. Vekua's theory of elliptic partial differential equations with analytic coefficients," *Zeitschrift für Angewandte Mathematik und Physik (ZAMP)*, vol. 8, pp. 169–203, 1957. 10.1007/BF01600500.

[10] S. Eisenstat, "On the rate of convergence of the Bergman–Vekua method for the numerical solution of elliptic boundary value problems," *SIAM Journal on Numerical Analysis*, vol. 11, no. 3, pp. 654–680, 1974.

[11] A. Moiola, R. Hiptmair, and I. Perugia, "Plane wave approximation of homogeneous Helmholtz solutions," *Zeitschrift für Angewandte Mathematik und Physik (ZAMP)*, vol. 62, pp. 809–837, 2011. 10.1007/s00033-011-0147-y.

[12] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inform. Theory*, vol. 50, pp. 2231–2242, 2004.

[13] G. Chardon, A. Cohen, and L. Daudet, "Reconstruction of solutions to the Helmholtz equation from punctual measurements," in *10th International Conference on Sampling Theory and Applications*, (Bremen, Germany), July 2013.

[14] H. Kuttruff, *Room acoustics*. Spon Press, fifth edition ed., 2009.

[15] M. Crocco, A. Del Bue, M. Bustreo, and V. Murino, "A closed form solution to the microphone position self-calibration problem," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 2597–2600, 2012.



**Gilles Chardon** received the engineering degrees of the École Polytechnique and Telecom ParisTech in 2009, as well as the MSc ATIAM of Université Pierre et Marie Curie, Paris VI. After working towards his PhD at Institut Langevin in Paris, and a postdoctoral position with the Mathematics and Signal Processing group of the Acoustics Research Institute of the Austrian Academy of Sciences in Vienna, he is now Associate Professor with CentraleSupélec, Gif-sur-Yvette, France. His main research interests include sparse representation of acoustical fields, inverse problems and numerical analysis in acoustics.



**Thibault Nowakowski** received a B.S. in science engineering from the University of Pierre et Marie Curie, France, and a B.A. in Musicoloy from the University of Sorbonne, France in 2010. He received a M.S. in acoustics from the University of Pierre et Marie Curie, France, in 2012. He is currently pursuing a Ph.D. on source localization in reverberant environments at the Institut Langevin (Paris, France).



**Julien de Rosny** received the M.Sc. Degree from the University Pierre et Marie Curie, Paris, France in 1996 and the Ph.D. degree from the same university in 2000, all in wave physics. During the 2000 – 2001 academic year, he was a postdoctoral fellow in the Marine Physical Laboratory at Scripps Research Institute, California, USA. In 2001, he became a junior scientist of CNRS (Centre National de la Recherche Scientifique) at Laboratoire Ondes et Acoustique, Paris, France. He's currently (since 2014) a CNRS senior scientist at Institut Langevin, Paris, France. His fields of interest include wave propagation and time reversal of waves in complex media, passive and active imaging applied to acoustic, elastic and electromagnetic waves.

Copyright (c) 2015 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JSTSP.2015.2422673





**Laurent Daudet** (M'04–SM'10) studied at the Ecole Normale Supérieure in Paris, where he graduated in statistical and non-linear physics. In 2000, he received a PhD in mathematical modeling from the Université de Provence, Marseille, France. After a Marie Curie post-doctoral fellowship at the C4DM, Queen Mary University of London, UK, he worked as associate professor at UPMC (Paris 6 University) in the Musical Acoustics Lab. He is now Professor at Paris Diderot University – Paris 7, with research at the Langevin Institute for Waves and Images, where he currently holds a joint position with the Institut Universitaire de France. Laurent Daudet is author or co-author of over 160 publications (journal papers or conference proceedings) on various aspects of acoustics and audio signal processing, in particular using sparse representations.

Copyright (c) 2015 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.