Acoustic sources joint localization and characterization using compressive sampling

Antoine Peillot, François Ollivier, Gilles Chardon, Laurent Daudet

ACOUSTICS 2012

Acoustic sources joint localization and characterization using compressive sampling

A. Peillot[a], F. Ollivier[a], G. Chardon[b] and L. Daudet[b]

[a]UPMC - Institut Jean Le Rond d'Alembert, 2 place de la gare de ceinture, 78210 Saint Cyr L'Ecole, France
[b]Institut Langevin, 10 rue Vauquelin, 75005 Paris, France
antoine.peillot@upmc.fr

In this work, a sparsity promoting strategy is developed in order to jointly achieve two complementary tasks regarding sound sources: localization and identification. Here, the sources are assumed sparse in the spatial domain, and greedy algorithms are used for their joint localization and identification in terms of spherical harmonics components. We study the case of sources located in a 3D space both numerically and experimentally. Two real sources based on loudspeaker elements are first precisely calibrated using a dense virtual spherical array. These standard sources are used afterwards to evaluate a sparse localization/identification strategy using an 3D array of microphones. Therefore signals from the calibrated sources are simulated over 96 randomly distributed microphones. The sparse model of the sources to be found mixes location and spherical harmonics decomposition. The identification strategy is evaluated according to a random positioning error of the microphones, and also according to the shift of the reconstruction grid with respect to the true position of the simulated sources. Finally the sparse identification strategy is implemented on signals recorded by random arrays of microphones in a real environment.

1 Introduction

Acoustic source identification is an important issue when dealing with noise problems. Identification aims at an objective description of the acoustic phenomenon generating a perturbing noise. A preliminary step to its identification is the precise localization of the source to be studied by processing the signals issued from an array of microphones. As a matter of fact this preliminary step is most of the time considered sufficient. Indeed space is generally investigated looking for simple monopoles, therefore obtaining no clues regarding the true nature of the sources. Anyway, identifying a source nature requires having a global look at it, which is generally not the case for standard techniques using microphones located in the far field of the source. The objective of the present work is to develop a strategy allowing performing in a 3D space the joint localization and objective characterization of sources of arbitrary nature. This strategy is based on the design of complex spherical harmonics (SH) modelling of the sources combined with the use of sparsity promoting algorithms for their localization. The second section of the paper presents the SH formalism used for describing complex directivities. The third section presents the numerical strategy using a greedy algorithm, Orthogonal Matching Pursuit (OMP), to achieve jointly the localization and identification tasks. This OMP algorithm relies on the careful design of a model 'dictionary' which is also described. The fourth section presents the experiment set up for calibrating arbitrary sources. A quantified identification of these sources is given in terms of the SH decomposition basis. These calibrated sources are used in the following for simulations and experiments. Simulations of the sparse analysis process are presented. In the fifth section we describe experiments set up for the identification of the calibrated sources using 3D random arrays of microphones. For the experiments led in a real environment the results provided by the previous strategy are discussed. And finally, concluding remarks expose its perspectives and limitations.

2 Source model

2.1 Spherical harmonics model

In order to have a physical comprehension of the sought sources, a natural way is to fit their radiation signature to models of known phenomena or a combination of them. In this framework we recall that a point mass flux is described by a monopole model, while a local momentum or force flux obeys a dipole model. These two models correspond to the zero and first order elements of the spherical harmonics family which is to be used in the following to decompose complex sources. Let an harmonic sound field spatially sampled in a 3-dimensional space be described by the spherical coordinates (O, r, Ω), where $\Omega \equiv (\theta, \phi)$. The spherical harmonic expansion of the sound pressure function p at any sample point (r, Ω), and wave number k, emanating from a point source at the origin writes, in free field conditions [1]

$$p(kr, \Omega) = \sum_{l=0}^{\infty} \sum_{q=-l}^{l} x_n^q(k) h_l(kr) Y_l^q(\Omega) \tag{1}$$

$Y_l^q(\cdot)$ denotes the spherical harmonic of order l and degree q, whose expression is given by:

$$Y_l^q(\Omega) \equiv \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-q)!}{(l+q)!}} P_l^q(\cos\theta)\, \mathrm{e}^{iq\phi} \tag{2}$$

where $P_l^q(\cdot)$ denotes the orthogonal Legendre polynomials and $h_l(\cdot)$ the corresponding outgoing radial Hankel function of order l. Both of these mathematical objects are defined in [1]. $x_l^q(k)$ are the spherical harmonics expansion coefficients.

In this work we intend to identify the radiation pattern of a complex source by estimating the unknown coefficients $x_l^q(k)$ of the spherical harmonics expansion. The pressure p is sampled with microphones at several positions in the sound field. Note that the considered pressures are harmonic at the specific frequency $f = kc/2\pi$, where c is the sound velocity.

However, a correct radiation analysis relies on the proper knowledge of the source location. Recent works have shown that translations and rotations of the source have significant effects on the expansion coefficients and may yield to highly inaccurate radiation pattern estimation [2, 3, 4]. Furthermore, the exact alignment of a source can be very difficult and even impossible to manage particularly for extended sources such as musical instruments. In the following, we propose a different method based on the joint localization of the acoustic center of a complex source and identification of the $x_l^q(k)$.

3 Joint localization and identification strategy

3.1 Problem formulation

The strategy consists in seeking the exact source location among N possible positions (r_n, Ω_n), nodes on a predefined

grid. Therefore, the radiated sound field is sampled at M points (r_m, Ω_m) using a set of microphones. For each possible target point (r_n, Ω_n), equation (1) can be rearranged in matrix form:

$$\mathbf{p} = \mathbf{A}_n \mathbf{x}_n \tag{3}$$

where $\mathbf{p}$, the $M \times 1$ measured pressure vector, writes:

$$\mathbf{p} = \left[p(r_1, \Omega_1) \ \dots \ p(r_m, \Omega_m) \ \dots \ p(r_M, \Omega_M)\right]^T \tag{4}$$

The dependence on k is omitted for notation simplicity. If the model is restricted to order L, we seek the $(L+1)^2 \times 1$ vector of SH expansion coefficients writing:

$$\mathbf{x}_n = \left[(x_0^0)_n \ (x_1^{-1})_n \ \dots \ (x_l^q)_n \ \dots \ (x_L^L)_n\right]^T \tag{5}$$

and the $M \times (L+1)^2$ transfer matrix $\mathbf{A}_n$ is defined by :

$$\mathbf{A}_n = [(\mathbf{a}_0^0)_n \ \dots \ (\mathbf{a}_l^q)_n \ \dots \ (\mathbf{a}_L^L)_n]^T \tag{6}$$

where

$$(\mathbf{a}_l^q)_n = \left[h_l(r_{1n}) Y_l^q(\Omega_{1n}) \ \dots \ h_l(r_{Mn}) Y_l^q(\Omega_{Mn})\right]^T \tag{7}$$

The $(\mathbf{a}_l^q)_{n=1\dots N}$ vectors are the elementary radiation patterns, also called 'atoms'. Now when dealing with the whole set of target points, equation (3) becomes:

$$\mathbf{p} = \mathbf{A}\mathbf{x} \tag{8}$$

where the $M \times N(L+1)^2$ matrix $\mathbf{A}$,also called 'dictionary', is the concatenation of the sub-arrays $\mathbf{A}_n$, so that :

$$\mathbf{A} = [\mathbf{A}_1 \ \dots \ \mathbf{A}_n \ \dots \ \mathbf{A}_N] \tag{9}$$

and the corresponding $N(L+1)^2 \times 1$ coefficients column vector $\mathbf{x} = [\mathbf{x}_1^T \dots \mathbf{x}_n^T \dots \mathbf{x}_N^T]^T$.

3.2 Sparsity promoting recovery

Solving the inverse formulation of system (8) yields the source vector $\mathbf{x}$ that contains the coefficients of the spherical harmonics expansion for each candidate position (r_n, Ω_n). The number of sampling points, i.e the number of microphones, is far lower than the dimension of the source vector: $dim(\mathbf{x})$. The system is consequently highly under-determined and has an infinity of solutions. Moreover, in our source localization framework, we assume that the number of point sources is small regarding $dim(\mathbf{x})$. The signal $\mathbf{x}$ is said to be sparse in the spatial domain. Let us recall that a signal $\mathbf{x}$ is said to be $K-$ sparse when it can be approximated using only K coefficients in a relevant decomposition basis, with $K \ll dim(\mathbf{x})$). Therefore $\mathbf{p}$ is a linear combination of the elements of only K sub-arrays $\mathbf{A}_j$ out of $\mathbf{A}$:

$$\mathbf{p} = \sum_{j \in J} \mathbf{x}_j \mathbf{A}_j \tag{10}$$

where J is a subset of $\{1 \dots N(L+1)^2\}$ such that $\mathrm{Card}(J) = K$. In this paper we focus on the selection of only one acoustic centre. The source vector $\mathbf{x}$ is therefore expected to be 1-sparse in the space domain. In parallel we do not consider sparsity in the elementary radiation patterns sub-domain.

Recent advances in signal processing have shown that a sufficiently sparse signal can be recovered accurately from a few measurements provided that the transfer matrix $\mathbf{A}$ presents specific properties. The theory, called Compressive Sampling, stipulates that if $\mathbf{A}$ presents some randomness, the signal $\mathbf{x}$ can be recovered from only $M = O(Klog(N/K))$ measurements [5]. Moreover, CS theoretically allows to under sample the measured field far below the Nyquist rate. And finally, randomly distributing in space a relatively small number of microphones should provide a well-adapted dictionary $\mathbf{A}$.

The recovery process itself uses dedicated algorithms that aim, from an infinite set of solutions, to extract the sparsest one.

3.3 Reconstruction algorithm

In order to recover the sparse vector $\mathbf{x}$ we use an algorithm inspired by the "greedy" algorithm: Orthogonal Matching Pursuit (OMP) [6]. It comprises two successive steps exploiting the same transfer matrix given by equation (9): the first step locates the source, the second step identifies its nature. In the first step, the measurement vector $\mathbf{p}$ is correlated to each atom of the dictionary $\mathbf{A}$. The most correlated atom indicates the source location and the subspace $\mathbf{A_n}$ describing the source. In the second step the selected subspace undergoes the following iterative process in order to derive the SH expansion coefficients in the $\mathbf{x_n}$ subspace :

1. Initialize a residue $r_0 = \mathbf{p}$ and an iteration counter $k = 0$
2. Extract the contribution of the SH expansion coefficient $(x_l^q)_n$ by projecting the residue r_k over the components $(\mathbf{a}_l^q)_n$ of the selected $\mathbf{A_n}$ such that:
$$(x_l^q)_n = \max_{(l,q)} \left| \left\langle (\mathbf{a}_l^q)_n, r_k \right\rangle \right|$$
3. Remove the orthogonal projection of r_k on every remaining and previously orthogonalized elements of the subspace $\mathbf{A_n}$
4. Iterate $k = k + 1$ until every SH expansion coefficients in the $\mathbf{x_n}$ subspace has been selected , that is up to $k = (L+1)^2$ when working with a L-limited order.

Note that the orthogonalization procedure of step 3 has a key role. In fact, the family composed of the $(\mathbf{a}_l^q)_n$ for one source position should constitute an orthogonal basis if the location n corresponds to the origin of the coordinate system and if the pressure field $\mathbf{p}$ is sampled on the unit sphere. With randomly distributed microphones, these conditions are not met. Since the components $(\mathbf{a}_l^q)_{l=1\dots L}$ are not orthogonal, when subtracting the projection of r_k over $\mathbf{a}_{l'}^{q'}$, the algorithm introduces new components in the directions of $\{\mathbf{a}_l^q\}_{l \neq l', q \neq q'}$. Therefore the components in $\mathbf{A_n}$ have to be orthogonalized. If several sources are to be located, the orthogonalization procedure has to be carried out for each selected sub-array. Interestingly it allows the use of different geometries of microphones arrays which may be more convenient to design than a spherical array.

4 Results

4.1 Source Calibration procedure

To evaluate our identification strategy, we need sources calibrated according to the SH decomposition basis. Two

sources based on standard loudspeaker elements have been chosen. The A source consists of a bare loudspeaker expected to approach a dipole type source. The B source is a baffled loudspeaker expected to approach a monopole. For their precise identification a set of 120 microphones was mounted according to a vertical semi-circle with a 110 cm radius. The elevation radiation pattern is collected with a 1.5 degrees step according to 36 azimuthal positions of the source. Since the calibration process is led in a real environment, it is necessary to get free of the reverberation field. Therefore a pulse compression scheme was adopted. The signal used to drive the source is a [50Hz-10kHz] linear chirp signal. The associated adaptive filtering provides a time resolution sufficient to evaluate properly the amplitude of the direct signal. As for the phase, is given by the sign of the compressed pulse, it equals 0 or π. The last calibration step consists in the identification of the HS components of the source according to the model described by equation (1). Figure 1 presents the measured directivity pattern of the bare loudspeaker (Calibrated A Source) and the baffled one (Calibrated B Source). Like in the following simulations and experiments, the analysis of the source is limited to order $L = 3$ *i.e.* 16 SH components.

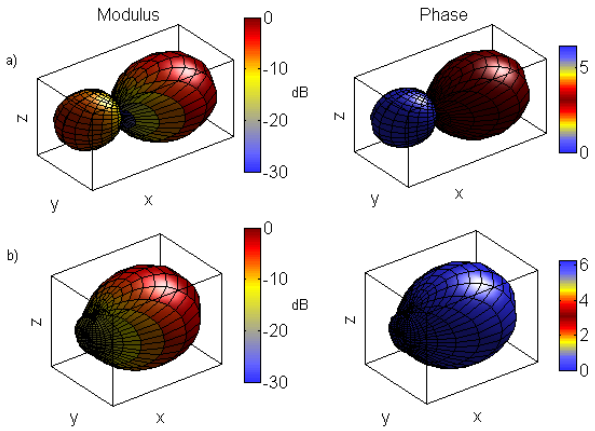

Figure 1: Measured directivity patterns for: a) a bare loudspeaker (Src A) and b) a baffled loudspeaker (Src B).

4.2 Numerical simulations

The sparse source localization/identification strategy was first tested in simulation scenarios. Two different 3D random microphone arrays surrounding the sources were tested : a spherical one (Array S) and an almost cubical one (Array C), more convenient to build for practical experiments. The accuracy of the source localization depends on the reconstruction grid step according to the wavelength λ. Since a large number of simulations has shown that the localization fails if the grid step $\Delta < \frac{\lambda}{2}$, we set it at $\Delta = \frac{\lambda}{2}$. We assume here that the actual source matches perfectly one of the reconstruction grid points. In our simulations the working frequency is 3kHz. For quantifying the quality of the sparse identification, we use a correlation coefficient defined by :

$$C = \frac{\tilde{\mathbf{x}}^T \mathbf{x}}{||\tilde{\mathbf{x}}||\,||\mathbf{x}||} \tag{11}$$

where $\tilde{\mathbf{x}}$ stands for the sparse estimation of $\mathbf{x}$.

4.2.1 Random spherical array

The calibrated sources were simulated to radiate in a free field, there complex radiation being captured by 96 microphones randomly distributed on the sphere identical to the calibrating one as shown on figure 2. In this case, the complex point source was located at the center of the sphere. In order to be representative of actual experimental conditions, the simulations involve a positioning error for the microphones as explained in section 4.2.3. This error is set at 8% of the wavelength to remain in the working domain of the process (cf. Figure 6). The result for the

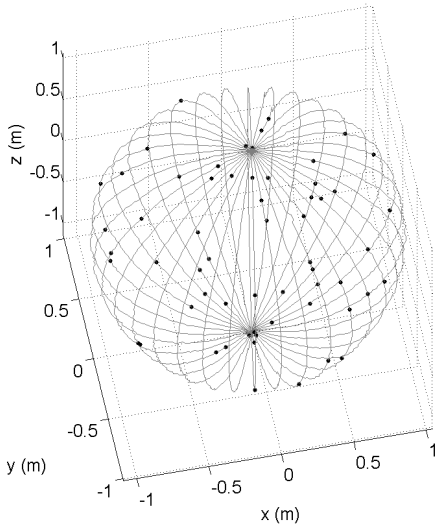

Figure 2: Random spherical array (Array S).

identification of the A source is shown on Figure 3. Apart an exact localization, it must be noticed that though apparently slightly distorted, the source is clearly identified in terms of SH components in amplitudes and phases as well. The identification reaches a 94% correlation coefficient.

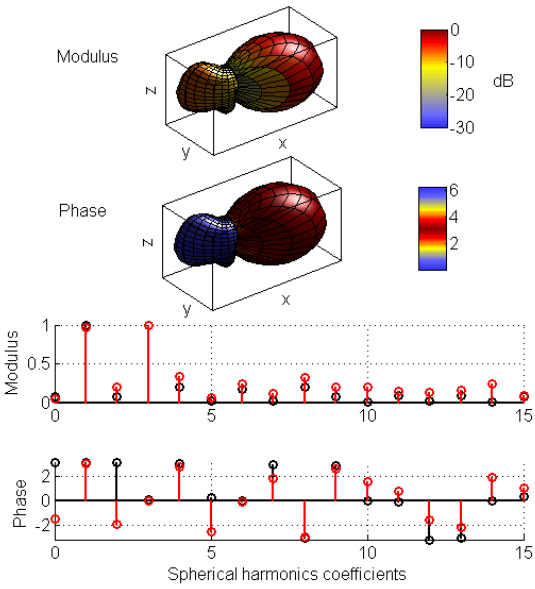

Figure 3: Identified directivity pattern and complex coefficients for the A source from signals simulated over array S. (Black : calibration - Red : sparse identification)

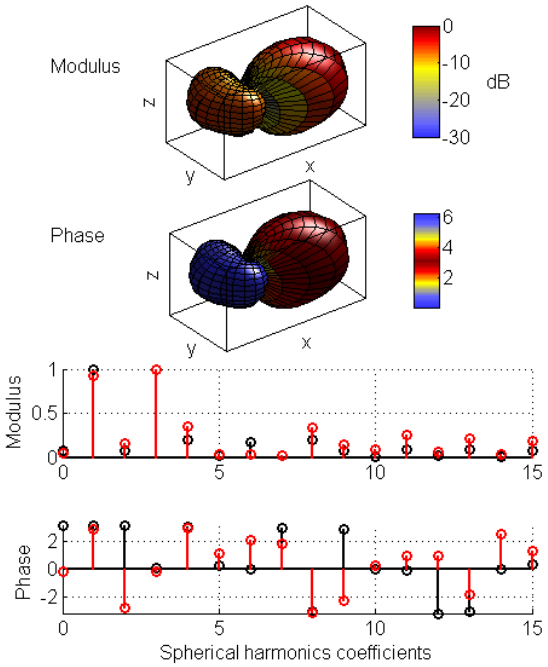

Figure 4: Identified directivity pattern and complex coefficients for source A from signals simulated over array C. (Black : calibration - Red : sparse identification)

4.2.2 Random arbitrary array

We considered secondly an almost cubical structure holding also 96 microphones randomly distributed on the edges of the structure (array C). Here the radial Hankel functions $h_l(kr)$ are of some importance since they comprise the microphones phase shifts. The identification result is shown on Figure 4. It shows to be very similar to the one obtained with array S and exhibits a correlation coefficient of 93% for the A source. The B source is also very precisely recovered (cf. Figure 5) with a correlation coefficient reaching 96%.

4.2.3 Sensitivity to sensors positioning errors and reconstruction grid shift

An important issue for a correct source recovery is the precise positioning of the numerous measuring microphones. This matter arises dramatically when dealing with experimental set ups. Unless a known rigid structure is available to hold the microphones, their precise location in a 3D-space is an awkward process. Therefore inevitable phase shifts of the measurement signals generally occur that may ruin the reconstruction process. Acoustic techniques such as the one described by Ono *et al.* [7] can help to overcome the problem partially, when the frequency is not too high. Nevertheless since exact positioning is hardly reachable in practice, the sensitivity of our strategy to microphones misplacement is to be investigated. Therefore the simulated signals were phase shifted according to a normal random law. The standard deviation of these phase shifts is evaluated with respect to λ. Figure 6 shows that microphone positioning errors causes increasing drop of the accuracy of the reconstruction but remain acceptable up to around a 8% of the wavelength. This means errors of the order of the centimetre at 3kHz, which is manageable.

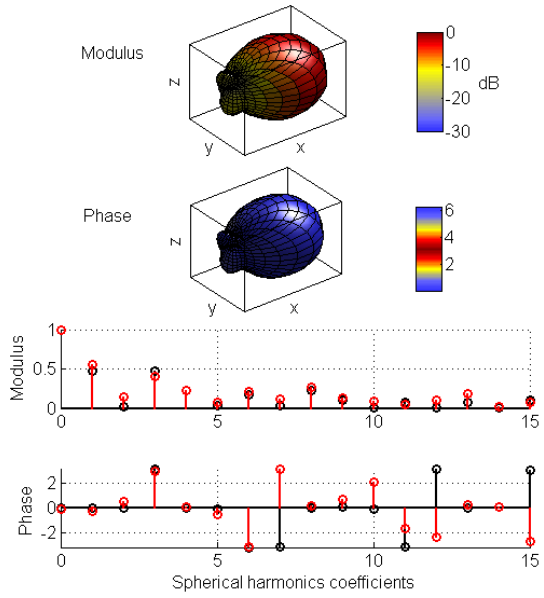

Figure 5: Identified directivity pattern and complex coefficients for Source B from signals simulated over array C. (Black : calibration - Red : sparse identification)

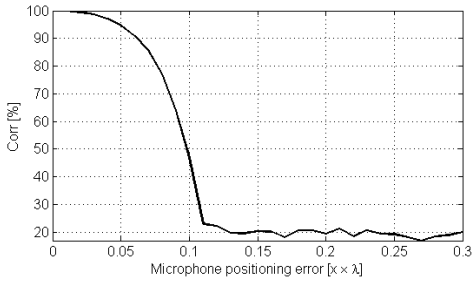

Figure 6: Averaged correlation results for source A between sparse identification using array S and calibration with respect to the random error in the microphone positioning.

Another problem of importance in our recovery process is the true location of the sought source with respect to the reconstruction grid nodes. This issue is similar to the source alignment problem discussed by Hagai *et al.* [3] or Pasqual *et al.* [4]. We investigated the sensitivity of our strategy to this shift. As shown on figure 7 the source recovery is acceptable up to a grid shift reaching approximatively a tenth of the wavelength.

4.3 Experiments

For the experiments, the source A was arbitrarily located at a point close to the approximative origin. We use the random spherical array S to recover the source. As shown on Figure 7 it is possible to identify the source with a good accuracy if the distance between the actual source and the closest grid point from the source is smaller than a tenth of the wavelength. On the other hand the reconstruction grid step has to be set at $\Delta = \frac{\lambda}{2}$. In order to jointly localize and identify the source we translate the grid 5 times in each direction and reiterate the process. Thus we obtain one solution for each position of the grid, i.e 125 solutions in the 3D coordinate system. An optimization criterion is needed to find the best approximation. The solution is inspired by

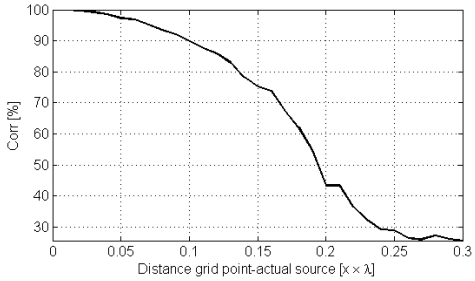

Figure 7: Averaged correlation results for source A between sparse identification using array S and calibration with respect to the shift of the reconstruction grid.

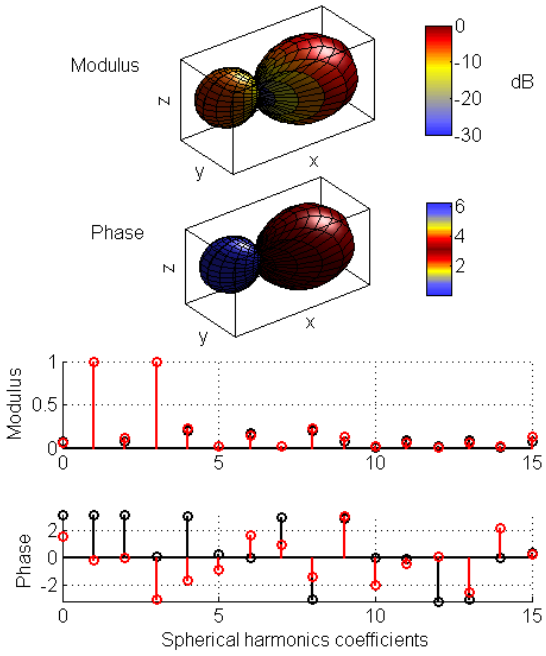

Figure 8: Identified directivity pattern and complex coefficients for Source A from signals recorded with array S. (Black : calibration - Red : sparse identification)

Rafaely [8] and his development of measures for source alignment. It uses the following cost function J for each position of the grid,

$$J = 1 - \frac{\alpha_{\tilde{L}}^2}{\|\mathbf{x}\|_2^2} \quad where \quad \alpha_{\tilde{L}}^2 = \sum_{q=-\tilde{L}}^{\tilde{L}} |x_{\tilde{L}}^q|^2$$

$\tilde{L}$ is the *a priori* predominant order of the source under study. Order 1 was chosen for source A. The solution of the identification problem is given by minimizing J over all positions of the grid. The result of the joint localization and identification of the source is shown on Figure 8. It exhibits a correlation coefficient of 91% which prove the efficiency of the method. Note that the phase components are all shifted from π, their relative value though are correct. However the source order is generally not known *a priori*. An objective optimization criteria is still under investigation.

5 Conclusion

The paper presents a new strategy to locate and identify complex acoustic sources in a 3-D space using random 3D-arrays of microphones. It takes advantage of a sparsity promoting algorithm for a precise localization and is based on a spherical harmonics analysis for the sources description. Simulations and experiments were carried out up to the third order of spherical harmonics. They have shown that the technique succeeds with a small number of microphones regarding the relatively high frequency of analysis ($3kHz$). However, while the strategy is efficient even with a reasonable misplacements of microphones, an optimization process for the choice of a reconstruction grid is under investigation. It could be included in the sparsity promoting algorithm for minimum computational costs. Finally the processing scheme is to be confronted to multiple acoustic centres scenarios.

Acknowledgments

This work was conducted in the framework of the ANR-ECHANGE project (ANR-08-EMER-006).

References

[1] E. G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*, Academic Press, London (1999)

[2] D. Deboy, Z. Zotter, "Acoustic center and orientation analysis of sound-radiation recorded with a surrounding spherical microphone array", *Proc. 2nd Int. Symp. ambisonics spherical acoustics*, 1-6, PAris, France (2010)

[3] I. B. Hagai, M. Pollow, M. Vorlander, B. Rafaely, "Acoustic centering of sources measured by surrounding spherical arrays", *J. Acoust. Soc. Am.* **130**(4), 2003-2015 (2011)

[4] A. M. Pasqual, V. Martin, "Inverse sound source reconstruction by exterior spherical acoustical holography with model adaptation", *Proc. 18th ICSV*, 1-8, Rio, Brazil (2011)

[5] E. Candes, M. B. Wakin, "An introduction to compressive sampling", *IEEE Signal Processing Magazine* **25**, 21-30 (2008)

[6] Y. C. Pati, R. Rezaiifar, Y. Rezaiifar and P. S. Krishnaprasad, "Orthogonal Matching Pursuit: Recursive function approximation with applications to wavelet decomposition", *27th Annual Asilomar Conference on Signals Systems and Computers* **130**(4), 40-44 (1993)

[7] N. Ono, H. Kohno, N. Ito, S. Sagayama, "Blind alignment of asynchronously recorded signals for distributed microphone array", *Proc. WASPAA*, 161-164, New Paltz, NY (2009)

[8] B. Rafaely, "Spatial alignment of acoustic sources based on spherical harmonics radiation analysis", *4th ISCCSP*, 1-5, Limassol, Cyprus (2010)