Block-sparse approach for the identification of complex sound sources in a room

Hugo Demontis, François Ollivier, Jacques Marchal

BLOCK-SPARSE APPROACH FOR THE IDENTIFICATION OF COMPLEX SOUND SOURCES IN A ROOM

Hugo Demontis[1] **Francois Ollivier**[1] **Jacques Marchal**[1]

[1] Sorbonne Université, CNRS, Institut Jean Le Rond dAlembert, UMR 7190, F-75005 Paris, France
hugo.demontis@sorbonne-universite.fr

ABSTRACT

This paper describes a practical methodology to estimate the directivity pattern of complex sound sources in reverberant environments. The acoustic analysis takes place under non-anechoic conditions. In this context, classical beamforming techniques fails to achieve this task because they generally rely on the free-field propagation model. In order to identify the directivity pattern with respect to frequency, an adapted block-sparse algorithm is proposed. This considers the first reflections to approximate the propagation path in the room, useful to solve accurately the related inverse problem. In addition, a large 3D array of microphones is implemented and validated. Largeness concerns here in both its dimensions and the number of microphones. The array consists of five planar sub-arrays which surround an entire room where sources are located. The microphones are flush mounted on the walls. From the pressure signal emitted by the source and recorded by the 1024 synchronous digital MEMS, a spherical harmonics representation of the source is then computed. A prototype of quadripolar source is tested to assert the efficiency of the proposed framework.

1. INTRODUCTION

From virtual reality to noise control, a wide range of applications in acoustics need the *directivity pattern* (DP) of a sound source to be known with a good approximation. To this end, surrounding spherical microphones arrays provide the most natural way to measure the DP experimentally [1]. In this type of antenna, sensors at fixed distance and distributed all around – i.e. at a solid angle of 4π – capture the direct sound field emitted by the complex sound source. For example, Behler proposes an arrangement of 32 microphones mounted on the vertices of a truncated icosahedron to evaluate the radiation nature of musical instruments. [2]. Spherical microphones array also improves the use of algorithms based on *spherical harmonics* (SH) formulation. The full 3-D pressure field representing the DP is then transformed to a 1-D vector containing few coefficients. The array design becomes here a relevant point, regarding the radius of the sphere, the number of microphones or their location along the spherical surface. This ensure to discard aliasing effects in the SH spectrum or to analyse the sound field along an extended frequency bandwidth. However, conventional recording systems generally suffer from low resolution in the SH domain, in part due to the small number of microphones used.

The DP estimation generally relies on a free field propagation model of the source. Most of the measurements are thus performed in an anechoic chamber. But under real conditions, like most of the time in confined spaces, specular reflections on the walls degrade the array signal. The performance of SH algorithms then decreases with the reverberation. In theory, considering the *Room Transfer Function* (RTF) to solve the inverse problem discard the room effects. But measuring a complete set of RTFs all over the volume is however unachievable due to the overwhelming number of samples needed to satisfy the Shannon sampling theorem [3].

The work in this paper presents the implementation of a complete methodology to overcome these two drawbacks. The numerical framework relies on the *Block Orthogonal Matching Pursuit* (BOMP) algorithm, which exploits the sparsity nature of the sources in the spatial domain [4]. The blocks refers in our case to the SH representation of the DP. This allows to joint the localization of the sources and their DP identification in a unique numerical task. The free-field version of the BOMP is then modified to be effective in reverberant environment, by including a closed form of the RTF based on the *Image-Source Method* (ISM) [5]. Only the early part is approximate here, involving the early specular reflections of the RTF. Beside we use a large 3D array with several hundred digital MEMS microphones. For practical reasons, the microphones are flush mounted on the walls and the ceiling of a rectangular room. Their location is known. Large aperture microphone arrays have already been studied in [6] and remains useful for sound source analysis along extended acoustic area.

This paper takes place in three parts. First, we establish the theory to understand the acoustic inverse problem regarding the estimation of the DP of sources. The BOMP algorithm is here introduced and its implementation is described. Secondly, we present the chosen acquisition system and the array deployment inside the room. Finally, we demonstrate the efficiency of the proposed method with an experiment. The SH nature of a quadripolar source, made with two unbaffled loudspeaker, is evaluated. Results are compared to the one obtained from previous work, involving a semicircular microphones array.

2. IDENTIFICATION FRAMEWORK

2.1 Source radiation model

Let a complex source at the origin of a 3D space described by the spherical coordinates $\mathbf{r} = (r, \Omega)$, with $\Omega = (\theta, \phi)$ i.e. the azimutal and zenithal couple. In free-field conditions, the continuous pressure distribution can be described at any point using the following SH expansion :

$$p(r, \Omega, k) = \sum_{l=0}^{L} \sum_{q=-l}^{l} \alpha_{lq}(k) h_l(kr) Y_l^q(\Omega) \qquad (1)$$

at the angular frequency $\omega = kc$, with wave number k and sound velocity c. The Hankel function of the second kind h_l refers to the radial acoustic spreading. The SH function Y_l^q of order l and degree q links with the angular dependency of the DP. More information about these mathematical objects can be found in [7]. The coefficients α_{lq} constitute the unknown in the DP estimation problem. They give the contribution of each SH to the pressure field and can be computed by the inverse transform of (1) :

$$\alpha_{lq}(k) = \frac{1}{h_l(kr)} \int_0^{2\pi} \int_0^{\pi} p(r, \Omega, k) Y_l^q(\Omega)^* d\Omega \qquad (2)$$

with $(\cdot)^*$ the conjugate operator. Note here that the pressure field is assumed to be band-limited up to the order L :

$$\alpha_{\tilde{L}q} = 0 \quad \forall \tilde{L} > L \qquad (3)$$

2.2 Discrete Formulation

The harmonic pressure field is sampled using M microphones. The signal p_m at the m^{th} microphone follows the free-field model in (1). The observations are stored in the vector $\mathbf{p}$ of size $M \times 1$. Sources are searched according to N grid points in a given volume of interest. For one at the n^{th} point, the matrix formulation of (2) writes :

$$\boldsymbol{\alpha}_n = \mathbf{H}_n^\dagger \mathbf{p} \qquad (4)$$

with $(\cdot)^\dagger$ the generalized inverse operator (or Moore-Penrose pseudoinverse). The set $\boldsymbol{\alpha}$ of size $(L+1)^2 \times 1$, containing the harmonic SH coefficients, is evaluated in a least-square sense. The steering matrix $\mathbf{H}_n$ of size $M \times (L+1)^2$ describes the propagation path from the n^{th} point-source to each microphone :

$$\mathbf{H}_n = [(\mathbf{h}_0^0)_n \ \dots \ (\mathbf{h}_l^q)_n \ \dots \ (\mathbf{h}_L^L)_n] \qquad (5)$$

The "atoms" $(\mathbf{h}_l^q)_n$

$$(\mathbf{h}_l^q)_n = [h_l(kr_{1n}) Y_l^q(\Omega_{1n}) \ \dots \ h_l(kr_{Mn}) Y_l^q(\Omega_{Mn})]^{\mathrm{T}} \qquad (6)$$

of size $M \times 1$ depends on the spherical coordinates (r_{mn}, Ω_{mn}) i.e. the location of the m^{th} microphone regarding the n^{th} point.

2.3 Block-sparse approach

For a number of grid points greater than the number of microphones ($M \ll N$), this discrete inverse problem is under-determined and an infinite number of solutions exists. However, it can be regularized by exploiting the sparsity nature of the sources in the spatial domain ($S \ll N$). This assumption yields to approximate $\mathbf{p}$ by a linear combination of only S blocks of $\mathbf{H} = \{\mathbf{H}_n\}_{n=1:N}$. In this case, greedy algorithms like the BOMP remains a useful method to solve it. While the algorithm dedicates to recover a signal with a small number of measurements, our goal here is to jointly localize and identify multiple sources in the same numerical scheme. The n^{th} selected block refers to the s^{th} source and its harmonic SH coefficients are then estimated.

The framework proposed in this work derivates from the BOMP. It comprises four successive steps exploiting the discrete formulation in (4). Let $\mathbf{r}^0 = \mathbf{p}$ and $\mathbf{S} = \emptyset$:

1. Selecte the most correlated block $\mathbf{H}_n$ with the pressure vector $\mathbf{p}$:

$$s = \arg\max_n \left(\|\mathbf{H}_n^\dagger \mathbf{r}^{(i-1)}\|_\infty \right)$$

2. Update the set J with the index j :

$$\mathbf{S}^{(i)} = \mathbf{S}^{(i-1)} \cup s$$

3. Compute the orthogonal projector $\mathbf{\Pi}_{J^{(i)}}$ built from the selected blocks :

$$\mathbf{\Pi}_{\mathbf{S}^{(i)}} = \mathbf{H}_{\mathbf{S}^{(i)}} \mathbf{H}_{\mathbf{S}^{(i)}}^\dagger$$

4. Update the residual $\mathbf{r}^{(i)}$:

$$\mathbf{r}^{(i)} = (\mathbb{I} - \mathbf{\Pi}_{J^{(i)}}) \mathbf{r}^{(i-1)}$$

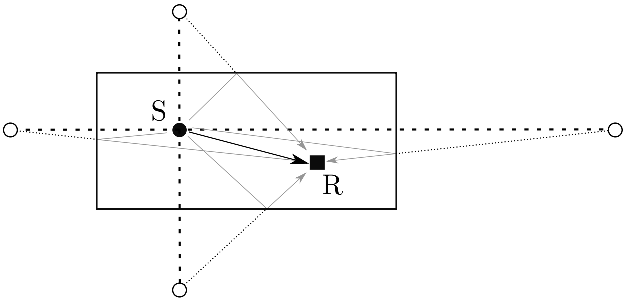
In case of S sources, these steps iterate until all the sources are discovered. Finally, the solution $\boldsymbol{\alpha}_\mathbf{S}$ is given as a matrix of size $(L+1)^2 \times S$:

$$\boldsymbol{\alpha}_\mathbf{S} = \mathbf{H}_\mathbf{S}^\dagger \mathbf{p} \qquad (7)$$

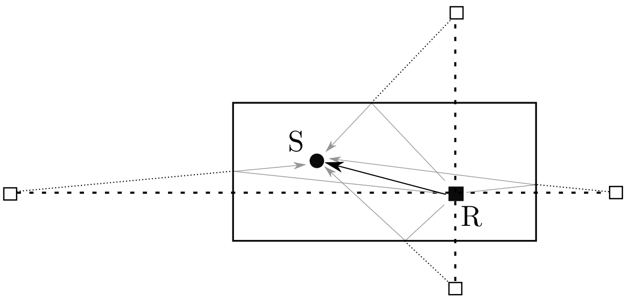
2.4 Under reverberant conditions

The steering matrix in (5) is only valid under anechoic conditions. Yet, algorithms based on free-field model tends to fail in rooms, when diffuse field interferes with the sources. This can be raised by including the RTF into the inverse problem solving. The RTF describes the propagation path between two points in an enclosed space. It takes into account the direct path and the multiple reflections caused by rigid boundaries and obstacles. For rooms perfectly rectangular, the ISM provides an efficient way to approximate it. The reflections are replaced by the free-field contributions from virtual sources (VS), located outside the walls. The real domain then extends to an infinity of virtual symmetric rooms. For each one mirrored, a virtual microphone array (VMA) can also be considered, recording the direct part

doi:10.25836/sasp.2019.17



(a) ISM



(b) IMM

Figure 1: Direct (ISM) and inverse (IMM) representation of the RTF.

of the associated VS. The total number J of VMAs relies directly with the order of reflections R. See [8] for a similar adaptation with classical delay-and-sum beamforming algorithm.

Applying the BOMP under reverberant conditions relies on this concept With the *Image-Microphone Method* (IMM), the process is performed not only with the real antenna, but also with all the considered VMAs. The implementation of the IMM stills straightforward. The computation of the parameters, like the reflection coefficients β_j or the spherical coordinates $(r^j_{mn}, \Omega^j_{mn})$ of the j^{th} VMA, is analog to those for the ISM.

2.5 Array description

The steering matrix in (5) is only valid under anechoic conditions. Yet, algorithms based on free-field model tends to fail in rooms, when diffuse field interferes with the sources. This can be raised by including the RTF into the inverse problem solving. The RTF describes the propagation path between two points in an enclosed space. It takes into account the direct path and the multiple reflections caused by rigid boundaries and obstacles. For rooms perfectly rectangular, the ISM provides an efficient way to approximate it. The reflections are replaced by the free-field contributions from virtual sources (VS), located outside the walls. The real domain then extends to an infinity of virtual symmetric rooms. For each one mirrored, a virtual microphone array (VMA) can also be considered, recording the direct part of the associated VS. The total number J of VMAs relies directly with the order of reflections R. See [8] for a similar adaptation with classical beamforming algorithm.

Applying the BOMP under reverberant conditions relies on this concept With the *Image-Microphone Method*

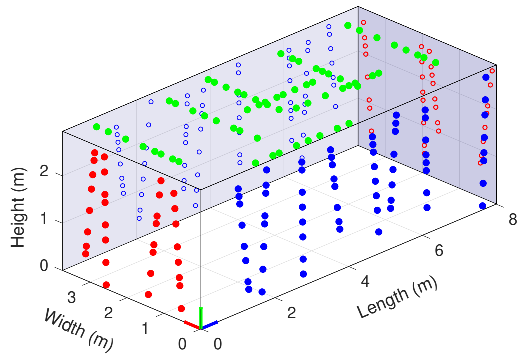


Figure 2: Apparent shape of the array. Only 256 nodes are plotted for visibility

(IMM), the process is performed not only with the real antenna, but also with all the considered VMAs. The implementation of the IMM stills straightforward. The computation of the parameters, like the reflection coefficients β_j or the spherical coordinates $(r^j_{mn}, \Omega^j_{mn})$ of the j^{th} VMA, is analog to those for the ISM.

3. EXPERIMENT

The acquisition system reposes on the use of digital MEMS microphones (InvenSense ICS43432). The integration of the conditioning circuitry and A/D conversion directly on the captor highly reduces the whole hardware complexity. Microphones are gathered by eight and plugged on a buffer to form beams. This buffers transmit the pressure datas to a sound card by RJ45 cables. Using USB3 protocole to communicate with computers, the plug-and-play interface allows the management of a huge numbers of sensors synchronously. More informations are provided in [9].

The array takes place in a rectangular room of $8.01m \times 3.75m \times 2.94m$. It comprises 1024 microphones directly flush-mounted to the walls and the ceiling in a pseudo-random distribution. The Figure 2 shows the in-situ proposed array . Only the floor is free of captors for more practicability during measurements. Its deployment induces errors in the position of the MEMS, compared to the predictions. These misalignments can affect the linear dependencies of the propagation model in (1). An acoustical geometric calibration permits to recover the effective geometry of the array. The steering matrix is then adjusted in the inversion in (7). We choose here the robust TOA-based algorithm in [10].

3.1 Source calibration

The asserting of our proposed framework requires to study sources according to the same SH basis of reference. A DP calibration procedure is thus performed before this experiment. The protocol employs a semicircular microphone array of radius $r = 1.19m$, similar to the one described in [11]. The pressure field from the tested source is captured by 104 MEMS microphones regularly spaced on the arc. A turntable rotates the source, to finally reproduce the spherical pressure distribution on a whole sphere. Due to the experimental setup, we choose to calibrate a controllable electroacoustic source. It brings together two naked loudspeakers, both facing in opposite directions with a spacing between them. Because of the theorical dipolar nature of each, the producing DP should be quadripolar. The corresponding SH vector $\boldsymbol{\alpha}_{ref}$ serves here as a reference. The term $(\cdot)_{expe}$ denotes results from the present experiment. The following correlation factor indicates the quality of the source identification :

$$\gamma = \frac{< \mathbf{p}_{ref}, \mathbf{p}_{expe} >}{\|\mathbf{p}_{ref}\|_2 \, \|\mathbf{p}_{expe}\|_2} \qquad (8)$$

The pressure vectors $\mathbf{p}_{ref}$ and $\mathbf{p}_{expe}$ are both reconstructed with $\boldsymbol{\alpha}_{ref}$ and $\boldsymbol{\alpha}_{expe}$ using (1).

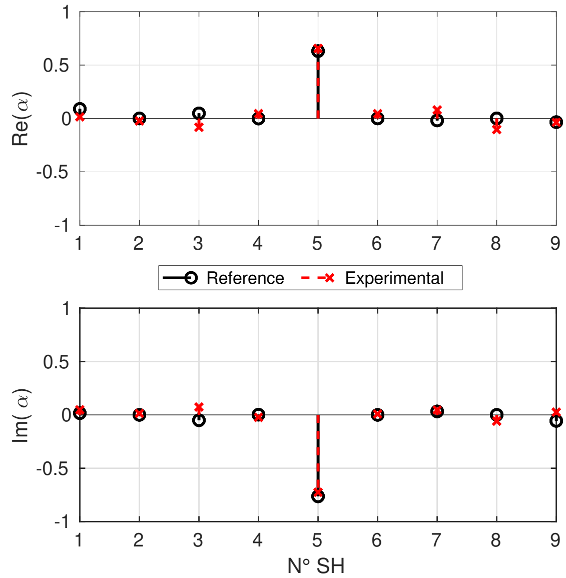
3.2 Results

The source locates now inside the room. The signal used to drive it is a [50Hz-10kHz] logarithmic chirp. The 1024 MEMS acquire synchronously the producing sound field. In addition, a reference microphone is placed inside the domain. Its frequency response and position are known.

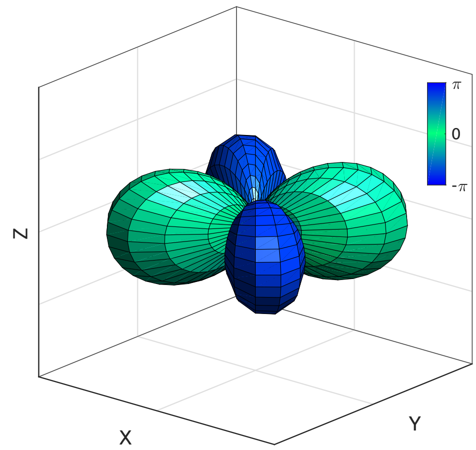
The Figure 3a shows the result of the present experiment, using the BOMP coupled with IMM, at frequency $f = 500$Hz. The reflection order is $R = 1$, generating 6 mirrors of the array. The maximum order sets to $L = 2$, describing the DP with 9 SH coefficients. The comparison between $\boldsymbol{\alpha}_{ref}$ and $\boldsymbol{\alpha}_{expe}$ is realized in term of real and imaginary parts. The x-axis represents the linear indexing of the SH basis ($l^2 + l + q + 1$). In both case, the same mode Y_2^{-2} is activated. The expected quadripolar behaviour is then accurately identified, as shown in Figure 3b. The correlation factor gives 96%, significating good agreements with the reference.

4. CONCLUSION

This paper describes a numerical procedure, validated by an experiment, in order to identify the directivity pattern of complex acoustic sources in real environments. An improvement of the BOMP is proposed for reverberant conditions by considering early reflections to solve the inverse problem. It remains straightforward and easy to implement. In parallel, an acquisition system comprising 1024 synchronous digital MEMS microphones is deploying in a typical rectangular room. The characterization of a quadripolar source is carried out. Results show good agreements with reference measurements and assert the efficiency of the proposed method. However, validations needs to be done for more scenarii, for example with many sources emitting at the same time or showing more complex radiativity. Uncontroled sound sources, like musical instruments, can also be investigated.



(a) SH decomposition of the source



(b) 3D view of the source

Figure 3: Results for the identification of the quadripolar source. The DP estimated in the present experiment $\boldsymbol{\alpha}_{expe}$ (in red) is compared to the reference $\boldsymbol{\alpha}_{ref}$ (in black).

5. REFERENCES

[1] M. Noisternig and F. Zotter, “On the decomposition of acoustic source radiation patterns measured with surrounding spherical microphone arrays,” *Proc. 37th DAGA*, 2011.

[2] G. Behler, M. Pollow, and M. Vorländer, “Measurements of musical instruments with surrounding spherical arrays,” in *Acoustics 2012*, 2012.

[3] T. Ajdler, L. Sbaiz, and M. Vetterli, “The plenacoustic function and its sampling,” *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3790–3804, 2006.

[4] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, “Block-sparse signals: Uncertainty relations and efficient recovery,” *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, 2010.

[5] R. Mignot, L. Daudet, and F. Ollivier, “Room reverberation reconstruction: Interpolation of the early part using compressed sensing,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 11, pp. 2301–2312, 2013.

doi:10.25836/sasp.2019.17

[6] H. F. Silverman and W. R. Patterson, "Visualizing the performance of large-aperture microphone arrays," in *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258)*, vol. 2, pp. 969–972, IEEE, 1999.

[7] E. G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*. Elsevier, 1999.

[8] A. Meyer, M. Pelz, and D. Dobler, "Microphone arrays an a wind tunnel environment with a hard reflective floor," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 258, pp. 3520–3530, Institute of Noise Control Engineering, 2018.

[9] C. Vanwynsberghe, R. Marchiano, F. Ollivier, P. Challande, H. Moingeon, and J. Marchal, "Design and implementation of a multi-octave-band audio camera for realtime diagnosis," *Applied Acoustics*, vol. 89, pp. 281–287, 2015.

[10] C. Vanwynsberghe, P. Challande, F. Ollivier, J. Marchal, and R. Marchiano, "Geometric calibration of very large microphone arrays in mismatched free field," *The Journal of the Acoustical Society of America*, vol. 145, no. 1, pp. 215–227, 2019.

[11] F. Ollivier, A. Peillot, G. Chardon, and L. Daudet, "Acoustic sources joint localization and characterization using compressed sensing," *The Journal of the Acoustical Society of America*, vol. 131, no. 4, pp. 3257–3257, 2012.