Sparse acoustic imaging with a spherical array

Efren Fernandez-Grande
Acoustic Technology, Department of Electrical Engineering, DTU - Technical University of Denmark, B. 352, DK-2800 Kgs. Lyngby, Denmark

Angeliki Xenaki
Department of Applied Mathematics and Computer Science, DTU - Technical University of Denmark, B. 303b, DK-2800 Kgs. Lyngby, Denmark

Summary

In recent years, a number of methods for sound source localization and sound field reconstruction with spherical microphone arrays have been proposed. These arrays have some useful properties, e.g. omni-directionality, robustness, compensable scattering, etc. This paper proposes a plane wave expansion method based on measurements with a spherical microphone array, solved in the framework provided by Compressive Sensing. The proposed methodology results in a sparse solution, i.e. few non-zero coefficients, and it is suitable for both source localization and sound field reconstruction. In general the method provides fine spatial resolution and is robust to noise (the noisy components are naturally suppressed). The validity and performance of the proposed method is examined, and some of its underlying assumptions are addressed.

PACS no. 43.60.+d, 43.58.+z

1. Introduction

Spherical microphone arrays are commonly used for analysing complex sound fields, e.g., localise sound sources in a given space [1–6], reconstruct the sound field near a source to examine how it injects energy into the medium [7–12], or capture complex acoustic scenes for subsequent reproduction with an array of loudspeakers [13, 14]. In recent years, spherical array processing is becoming increasingly popular for room acoustic applications. Unlike other common configurations [15], spherical arrays are omnidirectional, and therefore particularly well-suited in situations where sound waves impinge on the array from multiple directions, e.g. in environmental noise measurements, or in enclosed spaces such as rooms, vehicle interiors, etc.

Spherical microphone arrays have extensively been used for sound source localization and direction-of-arrival estimation. Recent studies have also examined the use of spherical arrays for the reconstruction of sound fields [7–12]; these methods use a spherical harmonic expansion to provide a representation of the sound field, used for extrapolating it to a different area than measured. A method was recently proposed that uses an elementary wave expansion to represent the measured sound field (using point sources), and solves for the corresponding coefficients via a matrix inversion, without explicit numerical integration on the sphere [11, 12].

This study proposes a method based on an elementary wave expansion that promotes a sparse solution to the problem (i.e. few non-zero coefficients), which aims at an optimal representation of the measured data, and results in greater spatial resolution and robustness to noise. The problem is formulated in the framework provided by Compressive Sensing (CS) [18]. Compressive Sensing makes use of the fact that signals that are sparse in some domain can be reconstructed perfectly, even with an apparently undersampled set of observations, by means of solving an l_1- minimization convex problem [17–19]. The proposed method consists of formulating an elementary wave expansion where the measured data is expanded in a basis of choice - plane waves in the present study, although also applicable to other wave functions. Then the problem is solved via l_1- minimization, instead of the conventional l_2- minimization (i.e., Least Squares), using convex optimization.

The focus of the present paper is to examine the proposed Compressive Sensing approach for spherical array processing, and discuss the potential benefits and limitations compared to the conventional least-squares approach.

2. Theory

2.1. Theoretical background

Let there be a rigid spherical array of radius a immersed in a sound field. The array consists of M microphones that are flush-mounted on the surface of the rigid-sphere. The sound pressure measured by the sensors can be expanded into plane waves arriving from every possible direction Ω_0,

$$p_t(a,\Omega_m) = \iint_{\Omega_0} -\frac{A(\Omega_0)}{(ka)^2} \times \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \frac{4\pi \mathrm{j}^{n+1}}{h_n'(ka)} Y_n^m(\Omega_m) Y_n^m(\Omega_0)^* d\Omega_0. \tag{1}$$

The summation over m and n corresponds to expanding a plane wave incident on a rigid sphere into spherical harmonics. The functions $h_n^{(2)}(x) = j_n(x) - \mathrm{j}y_n(x)$, are the spherical Hankel functions of the second kind (note that the sign convention chosen is $\mathrm{e}^{\mathrm{j}\omega t}$ with j the imaginary unit), and $j_n(x)$ and $y_n(x)$ are the spherical Bessel functions of the first and second kind [20]. Their derivative is represented by $j_n'(x)$ and $y_n'(x)$. We express the angular dependency as $\Omega \equiv (\theta,\phi)$, and $d\Omega \equiv \sin\theta d\theta d\phi$, so that the integration over the sphere is $\iint_\Omega (\cdot) d\Omega \equiv \int_0^{2\pi}\int_0^{\pi}(\cdot)\sin\theta d\theta d\phi$. Lastly, $Y_n^m(\Omega)$ are the spherical harmonics defined as in Ref. [20].

Equation 1 can be discretised and expressed as a sum of L plane waves (instead of the continuous expansion expressed by the integration over Ω_0)

$$p_t(a,\Omega_m) = \sum_{l=1}^{L} -\frac{A_l}{(ka)^2} \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \frac{4\pi \mathrm{j}^{n+1}}{h_n'(ka)} \times Y_n^m(\Omega_m) Y_n^m(\Omega_0)^*. \tag{2}$$

Alternatively, the wave expansion can be formulated in terms of point sources instead of plane waves. The measured pressure is expanded into a continuum of sources distributed over positions $\mathbf{r}_0$, associated with the integration surface S that can be chosen arbitrarily (it does not need to be spherical or even separable)

$$p_t(a,\Omega_m) = \iint_{S_0} -\frac{\mathrm{j}\rho c U(\mathbf{r}_0)}{a^2} \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \frac{h_n(kr_0)}{h_n'(ka)} \times Y_n^m(\Omega_m) Y_n^m(\Omega_0)^* dS, \tag{3}$$

where U is the surface velocity on S. In practice, the surface S can be placed either inside the source under study, or in general outside the domain in which the sound field is reconstructed, to prevent the singluarities. If the distribution of sources is discretised,

$$p_t(a,\Omega_m) = \sum_{l=1}^{L} -\frac{\mathrm{j}\rho c Q_l}{a^2} \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \frac{h_n(kr_0)}{h_n'(ka)} \times Y_n^m(\Omega_m) Y_n^m(\Omega_0)^*. \tag{4}$$

This approach based on a point source expansion (Eq. 4), is appropriate for the reconstruction of sound fields, because the decay of the acoustic field is modeled via the spherical spreading of the point sources. Contrarily, the former approach using plane waves (Eq. 2), is better suited for sound source localization where no extrapolation of the sound field is involved, but just aims at determining the direction from which sound waves impinge on the array, or sound field analysis. This is due to the fact that a (propagating) plane wave expansion basis cannot model decaying functions in a straightforward manner; evanescent waves should then be included.

2.2. Method

It is possible to express Eq. (2), i.e. the total sound pressure on the sphere, in matrix form, by conducting the summation over n and m

$$\mathbf{p} = \mathbf{H}\mathbf{x}, \tag{5}$$

noting that the summation should be truncated at N, to satisfy $ka < N$, depending on the number of microphones and size of the sphere [9]. The vector $\mathbf{p} \in \mathbb{C}^M$ contains the sound pressure measured at a discrete set of M points on the sphere, and the matrix $\mathbf{H}$ of dimensions $M \times L$ is the transfer matrix between the amplitude of the waves and the measured pressure. The amplitude of the waves corresponds to the vector $\mathbf{x} \in \mathbb{C}^L$, i.e., the unknown coefficients of the expansion. This problem is ill-posed, and most often underdetermined ($M < L$). Therefore the solution to Eq. (5) needs to be calculated via a regularized inversion of the problem.

In a general sense, the problem formulated in Eq. (5) can be cast as an optimization problem of the form

$$\min_{\mathbf{x}} ||\mathbf{x}||_p \quad \text{subject to } \mathbf{p}_t = \mathbf{H}\mathbf{x} \tag{6}$$

where $||\cdot||_p$ represents the vector norm,

$$||\mathbf{x}||_p = \left(\sum_i x_i^p\right)^{1/p}. \tag{7}$$

The choice of p-norm in the coefficient vector will promote different solutions to the problem.

The l_2-norm leads to the 'conventional' least-squares minimisation problem. This problem has the well-known closed form analytical solution (including a regularisation term λ),

$$\mathbf{x} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \lambda\mathbf{I}\right)^{-1} \mathbf{p}. \tag{8}$$

which has many non-zero components. The superscript H denotes the Heremitian transpose or conjugate transpose, $\mathbf{I}$ is the identity matrix and λ the regularization parameter for the Tikhonov regularization [10, 21]. When the regularization parameter is zero, the inversion corresponds to a matrix pseudo-inverse.

On the other hand, the choice of the l_0-pseudo norm in Eq. (6), defined as $||\mathbf{x}||_0 = \#(i|x_i \neq 0)$, promotes sparse solutions since, by definition, it minimises the number of non-zero entries in the vector. This norm is in fact a count of the non-zero elements of the vector. However, the l_0-norm minimisation is a combinatorial problem which often becomes intractable.

It can be shown [17] that for sufficiently sparse problems, the l_0-norm minimisation problem is equivalent to the l_1-norm problem. The choice of the l_1-norm promotes sparsity on the solution, i.e., a small number of non-zero coefficients, while it leads to a convex optimisation problem that can be solved efficiently.

In the presence of noise, the l_1 minimisation problem is reformulated as,

$$\min_{\mathbf{x}} ||\mathbf{x}||_1 \quad \text{subject to } ||\mathbf{H}\mathbf{x} - \mathbf{p}||_2 \leq \varepsilon. \qquad (9)$$

where ε is the noise floor. Alternatively, the problem can be formulated in an unconstrained form by introducing a regularisation parameter λ which determines the weight of the l_1-norm penalty.

One important aspect of the l_1 minimization solution is the incoherence between the columns of the sensing matrix $\mathbf{H}$, since this is a guarantee of the equivalence between the l_0 and l_1 minimization problems. The coherence between the columns of $\mathbf{H}$ can be described by the maximum of the non-diagonal elements of the Gram matrix

$$\mu(\mathbf{H}) = \max_{\{i \neq j\}} |\mathbf{H}^H \mathbf{H}|. \qquad (10)$$

This is also connected to the restricted isometry property (RIP) condition, $(1 - \delta_s)||\mathbf{x}||_2^2 \leq ||\mathbf{H}\mathbf{x}||_2^2 \leq (1 + \delta_s)||\mathbf{x}||_2^2$, which is frequently found in the literature.

The obtained vector $\tilde{\mathbf{x}}$ yields the amplitudes of the waves used in the expansion, and thus it is sufficient for estimating the direction of arrival of the waves. Nonetheless, if the aim were to reconstruct the sound-field elsewhere than measured, it is possible to extrapolate the wave expansion to reconstruct/predict the sound-field elsewhere: a reconstruction matrix $\mathbf{H}_s$ is defined (analogous to the one in Eq. 2 or 4) that relates the obtained expansion coefficients $\tilde{\mathbf{x}}$ to the desired reconstruction points in the medium $\mathbf{r}_s$ $\mathbf{p}_s = \mathbf{H}_s\, \tilde{\mathbf{x}}$, with the reconstructed pressure $\mathbf{p}_s \in \mathbb{C}^K$, and using a vector $\mathbf{r}_s$ instead of the $\mathbf{r}_0$ in Eq. (4). Alternatively, one can choose a reconstruction matrix based on the propagation of plane waves or point sources in free-space, thus compensate for the scattering of the array, and reconstruct the incident pressure as if the array was not present - see Ref. [12]).

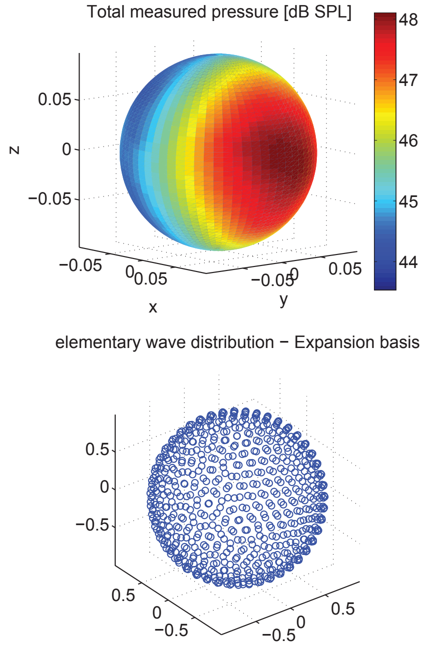

Figure 1. Top: Sound pressure on the spherical array; Bottom: Plane wave model: Distribution of waves used to model the soundfield - each circle corresponds to a direction (total of 650 waves).

3. Results

The proposed methodology and its numerical properties are examined through a simulation example. The method consists of the plane-wave expansion described in Eqs. (1) (2) and (5); the aim is to examine the solution obtained with the proposed CS methodology (Eq. 9), and to compare it with the conventional least-squares solution (Eq. 8).

The study consists of a point source located 6 m away from the array surface, a rigid spherical array of 50 microphones, near-uniformly distributed over its surface, that can sample up to 5 orders of spherical harmonics [9]. The pressure on the array surface at 500 Hz is shown in Fig. 1 (top). Normally distributed noise is added to the simulated measurements with a signal-to-noise ratio (SNR) of 25 dB. A plane wave expansion of 650 waves is considered, and shown in Fig. 1 (bottom). The waves are distributed over an equal solid angle spacing; the distribution of the waves has been determined based on a Thomson problem, that considers equally charged particles on a sphere, and are therefore uniformly distributed over the surface of the sphere.

Figure 2 shows the solution of the problem based on the conventional least-squares solution Eq. (8), and Figure 3 shows the proposed compressive sensing solution based on Eq. (9). The top row of both figures shows the recovered pressures by each of the techniques, making it apparent that both approaches can successfully explain the measured data. The center

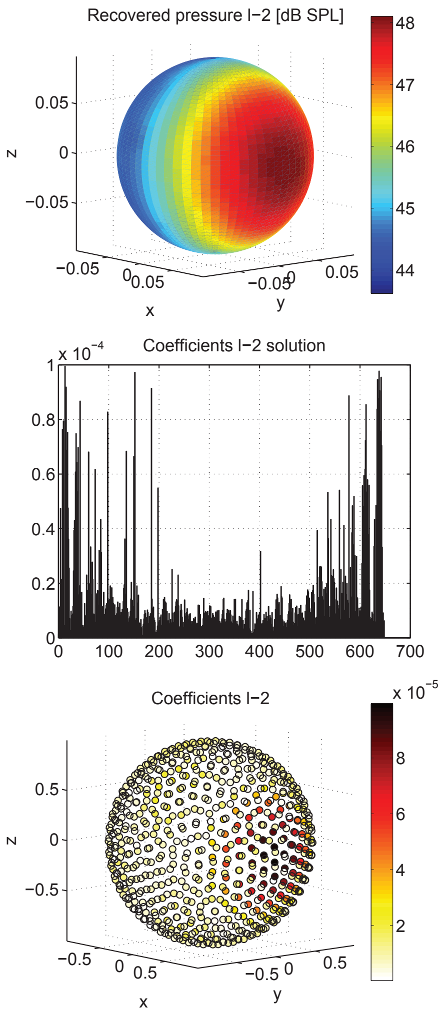

Figure 2. Spherical wave incident on a spherical array, expanded into plane waves (Eq. 2). Solution based on the l_2-norm - least squares (Eq. 8); Top: Recovered sound pressure; Mid: coefficients based on l_2 minimization; Bottom: Coefficients ordered by their direction of arrival.

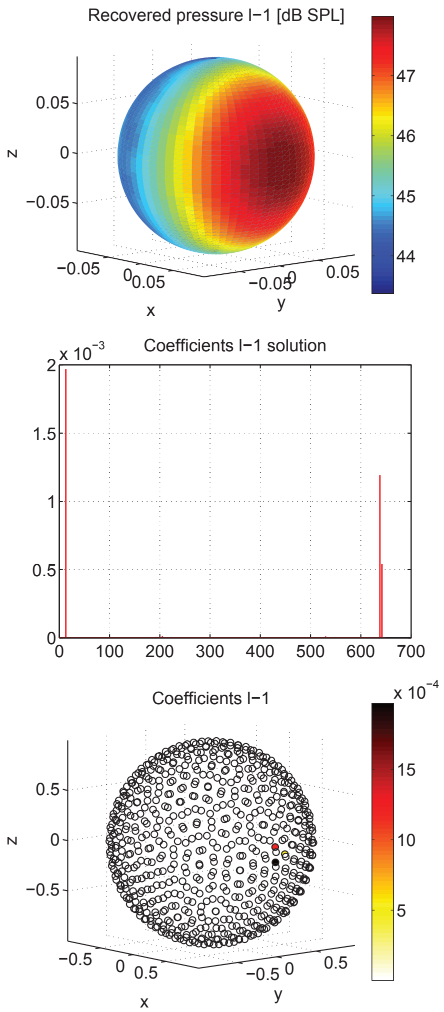

Figure 3. Spherical wave incident on a spherical array, expanded into plane waves (Eq. 2). Solution based on the l_1-norm - Compressive Sensing (Eq. 9); Top: Recovered sound pressure; Mid: coefficients based on l_1 minimization; Bottom: Coefficients ordered by their direction of arrival.

row shows the coefficients of the two methodologies (without any particular ordering), showing that the set of coefficients obtained by the two approaches is significantly different, particularly regarding the sparsity of the solution. In the l_2 least-squares approach (Fig. 2, mid.) all of the coefficients are non zero, whereas the proposed CS method returns only three non-zero coefficients (Fig. 3, mid.), indicating that the measured data can be explained optimally with a minimal number of entries in the expansion. The bottom plots of the figures show the coefficients of the two approaches and the corresponding direction of incidence of the plane waves used in the model (i.e. incoming direction of the waves used in the expansion). The least squares solution corresponds to a conventional beamformer output [4, 5], with its characteristic limited spatial resolution (main-lobe and side-lobes). The CS solution essentially uses three coefficients, detecting the direction of arrival of the spherical wave, where the source is located, with optimal accuracy.

All in all, it is apparent that although the recovered sound pressure by the two methods is virtually identical, the differences in the coefficients are notorious. Several aspects to consider about the Compressive Sensing approach / l_1- solution: i) the spatial resolution is enhanced and approaches an ideal delta function, ii) noise is naturally suppressed; it is evident how the sparsity constrain acts as a regularization mechanism robust to noise, since the recovered coefficients are those that span the signal, and not the measurement noise; iii) lastly, the resulting solution is based on only a few terms of the expansion, which also has the useful implication that an over-determined system of equations can be solved with only the relevant expansion terms, for better quantitative accuracy.

This study focuses on the description of the proposed methodology, and examines a single source case to provide an initial insight on its performance, and some of the numerical properties of the solution. Further studies will focus on sensitivity to noise, multiple sources and complex sound fields, as well as sources of larger spatial extent.

4. CONCLUSIONS

This study proposes and examines a method that makes use of the framework provided by Compressive Sensing for spherical array processing. The solution is obtained via l_1- minimization, which imposes sparsity on the solution, i.e. requires that few coefficients are non-zero. The results show that the method makes it possible to identify the direction of arrival of the waves with almost ideal spatial resolution and accuracy. This is the case provided that the coherence of the matrix columns is sufficiently low. In a general sense, the results indicate that it is possible to accurately represent the measured data with minimum number of coefficients. This gives the method a promising perspective for its use on direction of arrival estimation, sound field reconstruction and sound field analysis.

Acknowledgement

This work was supported by the Danish Council for Independent Research (DFF), under the individual postdoctoral grant FTP/12-126364.

References

[1] Jens Meyer and Gary Elko. A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. In *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, volume 2, pages II–1781–II–1784, May 2002.

[2] B Rafaely. Analysis and design of spherical microphone arrays. *Speech and Audio Processing, IEEE Transactions on*, 13(1):135–143, 2005. ISSN 1063-6676.

[3] Munhum Park and Boaz Rafaely. Sound-field analysis by plane-wave decomposition using spherical microphone array. *J. Acoust. Soc. Am*, 118(5):3094–3103, 2005.

[4] Karim Haddad and Jo/rgen Hald. 3d localization of acoustic sources with a spherical array. *Acoustics '08 Paris. Vol 123*, 123(5):3311–3311, 2008.

[5] Elisabet Tiana-Roig, Finn Jacobsen, and Efren Fernandez-Grande. Beamforming with a circular array of microphones mounted on a rigid sphere (L). *J. Acoust. Soc. Am*, 130(3):1095–1098, 2011.

[6] A. Pereira and Q. Leclere. Improving the equivalent source method for noise source identification in encolsed spaces. *Proceedings of ICSV 18, Rio de Janeiro, Brazil, 10-14 July 2011*, art. 055010, 2011.

[7] E.G. Williams, N. Valdivia, and P. C. Herdic. Volumetric acoustic vector intensity imager. *J. Acoust. Soc. Am.*, 120(4):1887–1897, 2006.

[8] E.G. Williams and K. Takashima. Vector intensity reconstructions in a volume surrounding a rigid spherical microphone array. *J. Acoust. Soc. Am.*, 127(2):773–783, 2010.

[9] Finn Jacobsen, Guillermo Moreno-Pescador, Efren Fernandez-Grande, and Jo/ rgen Hald. Near field acoustic holography with microphones on a rigid sphere (l). *The Journal of the Acoustical Society of America*, 129(6):3461–3464, 2011.

[10] Alba Granados, Finn Jacobsen, and Efren Fernandez-Grande. Regularised reconstruction of sound fields with a spherical microphone array. *Proceedings of Meetings on Acoustics*, 19(1): 055010, 2013.

[11] Efren Fernandez-Grande. Reconstruction of arbitrary sound fields with a rigid-sphere microphone array. *166th Acoust. Soc. Am. meeting San Francisco, 2-6th December*, 134(5):3998–3998, 2013.

[12] Efren Fernandez-Grande and Tim Walton. Reconstruction of sound fields with a spherical microphone array. In *Proceedings of Inter-noise 2014*, 2014.

[13] T.D. Abhayapala and Darren B. Ward. Theory and design of high order sound field microphones using spherical microphone array. In *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, volume 2, pages II–1949–II–1952, May 2002.

[14] Marton Marschall and Jiho Chang. Sound-field reconstruction performance of a mixed-order ambisonics microphone array. *Proceedings of Meetings on Acoustics*, 19(1):055007, 2013.

[15] Efren Fernandez-Grande, Finn Jacobsen, and Quentin Leclere. Sound field separation with sound pressure and particle velocity measurements. *The Journal of the Acoustical Society of America*, 132(6):3818–3825, 2012.

[16] B. Rafaely, B. Weiss, and E. Bachmat. Spatial aliasing in spherical microphone arrays. *Signal Processing, IEEE Transactions on*, 55(3):1003–1010, March 2007.

[17] Michael. Elad. *Sparse and redundant representations : From theory to applications in signal and image processing.* Springer, New York, Ch. 1, 2010. ISBN 9781441970107.

[18] E. J. Candes and M. B. Wakin. An introduction to compressive sampling. *IEE Signal Proc. Mag.*, 25: 21–30, 2008.

[19] Angeliki Xenaki, Peter Gerstoft, and Klaus Mosegaard. Compressive beamforming. *The*

Journal of the Acoustical Society of America, 136 (1):260–271, 2014.

[20] Earl G. Williams. *Fourier Acoustics - Sound Radiation and Nearfield Acoustic Holography.* Academic Press, 1999.

[21] Per Christian Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion.* SIAM monographs on mathematical modeling and computation. SIAM, 1998. ISBN 9780898714036.