Compressive sensing with a spherical microphone array

Fernandez Grande, Efren; Xenaki, Angeliki

Published in:
Journal of the Acoustical Society of America

Link to article, DOI:
10.1121/1.4942546

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Compressive sensing with a spherical microphone array

Efren Fernandez-Grande and Angeliki Xenaki

Acoustic Technology, Department of Electrical Engineering, Technical University of Denmark (DTU), Building 352, Ørsteds Plads, DK-2800 Kongens Lyngby, Denmark
efg@elektro.dtu.dk, anxe@elektro.dtu.dk

Abstract: A wave expansion method is proposed in this work, based on measurements with a spherical microphone array, and formulated in the framework provided by Compressive Sensing. The method promotes sparse solutions via $\ell_1$-norm minimization, so that the measured data are represented by few basis functions. This results in fine spatial resolution and accuracy. This publication covers the theoretical background of the method, including experimental results that illustrate some of the fundamental differences with the “conventional” least-squares approach. The proposed methodology is relevant for source localization, sound field reconstruction, and sound field analysis.

© 2016 Acoustical Society of America

Date Received: June 16, 2015 Date Accepted: January 27, 2016

1. Introduction
Spherical microphone arrays are well-suited for the analysis of complex sound fields, particularly in situations where sound waves impinge on the array from multiple directions (e.g., room acoustics, vehicle enclosures, environmental noise measurements, etc.). Due to their geometry and directional properties, spherical arrays are commonly used in these problems to localize sound sources,1–5 reconstruct the sound field near a source to examine how it radiates sound,6–11 or capture spatial features of a sound field for its reproduction with a loudspeaker array.12

The existing methods used for these problems typically rely on solving an underdetermined system of linear equations to obtain the amplitude of the waves that impinge on the array, or alternatively, to obtain the coefficients of a wave expansion used to represent the data captured in the measurement. The solution to this problem is usually sought in a least squares (LS) sense, which is not necessarily the most suitable choice (in fact, the choice can be justified by its analytical and numerical convenience, but not always by its physical significance). An alternative approach based on the framework provided by Compressive Sensing (CS) is proposed in this study. CS makes use of the fact that signals that are sparse in some domain can be reconstructed perfectly, even with an apparently under-sampled set of observations, by means of solving an $\ell_1$-minimization convex problem.13–15 The CS methodology promotes sparsity, i.e., few non-zero coefficients in the obtained solution.

The method presented in this study seeks a representation of the measured data using an elementary wave expansion, hence the data measured with the spherical microphone array are expanded into plane waves, and the resulting system of linear equations is solved in the CS framework via $\ell_1$-minimization—instead of the conventional LS solution.11 This yields an optimal solution with as few non-zero coefficients as possible. The proposed method results in high spatial resolution and accuracy.

2. Theory

2.1 Theoretical background
Let us consider a rigid spherical array that is immersed in a sound field, consisting of $M$ microphones flush-mounted on the surface of the rigid-sphere, of radius $a$. The measured sound pressure can be expanded into a continuum of plane waves arriving from every direction $\Omega_0$,

$$p(a, \Omega_m) = \int_{\Omega_0} \frac{A(\Omega_0)}{(ka)^2} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} 4\pi(-1)^{n+1} \frac{Y_n^m(\Omega_m) Y_n^m(\Omega_0)^*}{\tilde{h}_n(ka)} d\Omega_0.$$  \hfill (1)

$^a$Author to whom correspondence should be addressed.
where use of the Wronskian relationship has been made to the problem. Selecting a different norm in the coefficient vector will promote different solutions. The angular dependency is expressed as

\[ I(r, \Omega) = \sum_{n=0}^{\infty} \sum_{m=1}^{L} A_l Y_m^m(\Omega) Y_n^m(\Omega_0), \]

where \( Y_n^m(\Omega) \) is the spherical Bessel function. The summation over \( n \) and \( m \) corresponds to expanding a plane wave incident on a rigid sphere into spherical harmonics. Each wave is incident from direction \( \Omega_0\equiv(\theta_0, \phi_0) \) with amplitude \( A(\Omega_0) \), and the microphone positions on the sphere are \( r_m=(a, \Omega_m) \). The origin of coordinates is in the center of the sphere. The functions \( Y_n^m(\Omega) \) are the spherical harmonics of order \( m \) and degree \( n \), defined as in Ref. 16, \( h_n^m(x) \) is the spherical Hankel function of the second kind with \( j_m(x) \) and \( y_m(x) \) the spherical Bessel functions of the first and second kind, and \( h_n^m \) is the derivative of the Hankel function. Note that the sign convention chosen is \( e^{i\theta} \).

The angular dependency is expressed as \( \Omega \equiv (\theta, \phi) \), and the integration over the sphere is \( \int d\Omega = \int_0^{2\pi} d\theta \sin \theta d\phi \). Equation (1) can be discretized into a sum of \( L \) plane waves

\[ p(a, \Omega_m) = \sum_{l=1}^{L} A_l \sum_{m=0}^{\infty} \sum_{n=-m}^{m} \frac{4\pi(-j)^{n+1}}{h_n'(ka)} Y_m^m(\Omega_m) Y_n^m(\Omega_0), \]

where use of the Wronskian relationship has been made \( j_n(k a h'_n(k a) - j'_n(k a) h_n(k a) = -j/(ka)^2 \).

The plane wave model in Eq. (2) is well-suited for far-field measurements and sound source localization, where no extrapolation of the sound field takes place, and the wavefronts can be regarded as locally planar. This is the model considered in the following. However, it should be noted that a different wave basis can be used, e.g., the wave expansion can be formulated in terms of point sources instead of plane waves. The measured pressure is expanded into a set of sources distributed over positions \( r_0=(r_0, \Omega_0) \), which in discrete form is formulated as

\[ p(a, \Omega_m) = \sum_{l=1}^{L} \left( -j \frac{\rho c Q_l}{a^2} \right) \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{h_0(k a)}{h'_n(k a)} (\Omega_m) Y_n^m(\Omega_0), \]

where \( Q_l \) is the volume velocity of each source. This expansion is suitable for sound field reconstruction or near-field processing, for the decay of the acoustic field is modeled via the spherical spreading of the point sources.

2.2 Method

By conducting the summation over \( n \) and \( m \) in Eq. (2), a transfer matrix \( H \) is obtained that relates the pressure on the sphere and the coefficients of the wave model

\[ p = H x. \]

The summation should be truncated at \( n=N \), given the number of microphones and size of the sphere, satisfying \( ka < N \), where \( k \) is the wave number in the medium. The vector \( p \in \mathbb{C}^M \) consists of the sound pressure measured at a discrete set of \( M \) points on the sphere and \( H \in \mathbb{C}^{M \times L} \), is the transfer matrix between the amplitude of the waves and the measured pressure. The amplitude of the waves corresponds to the vector \( x \in \mathbb{C}^L \), i.e., the unknown coefficients \( A_l \) of the expansion in Eq. (2). This problem is ill-posed, typically underdetermined \( (M < L) \), and requires regularization. In general sense, the problem in Eq. (3) can be solved as an optimization problem of the form

\[ \min_x \|x\|_p \quad \text{subject to} \quad p = H x, \]

where \( \| \cdot \|_p \) is the \( \ell_p \) vector norm,

\[ \|x\|_p = \left( \sum_i |x_i|^p \right)^{1/p}. \]

The choice of different norms in the coefficient vector will promote different solutions to the problem. Selecting a \( \ell_2 \)-norm leads to a “conventional” LS minimization problem, which has the closed-form analytical solution

\[ x = H^H (HH^H + \lambda I)^{-1} p, \]

where \( I \) is the identity matrix, \( \lambda \) is the regularization parameter for the Tikhonov regularization, and the superscript \( H \) denotes the Hermitian transpose. When the regularization parameter is zero, the inversion corresponds to a conventional matrix pseudo-inverse. The LS solution tends to yield many non-zero components. In the case of sparse problems (as often found in acoustic array processing, i.e., few sources or waves impinging on the array, or few modes of an expansion) the LS solution provides a poor representation of the measured data, compromising the spatial resolution and dynamic range.
An interesting alternative is to choose a $\ell_0$ pseudo-norm in Eq. (4), defined as a count of the non-zero elements of a vector $\|x\|_0 = \#$ ($l \neq 0$), since it will promote sparse solutions. However, the $\ell_0$-norm minimization is a combinatorial problem that often becomes intractable, thus it is not a feasible approach. CS postulates that for sparse problems, the solution to the $\ell_0$-norm minimization problem can be equivalent to the $\ell_1$-norm minimization one, and importantly, the $\ell_1$-norm leads to a convex optimization problem that can be solved efficiently. Consequently the choice of a $\ell_1$-norm will promote a sparse solution (i.e., an optimal representation of the measured data with minimal number of non-zero terms). The $\ell_1$-minimization problem can be formulated as

$$\min_{x} \|x\|_1 \quad \text{subject to} \quad \|Hx - p\|_2 \leq \varepsilon,$$

where $\varepsilon$ is the noise floor present in the measurement vector. We make use of Eq. (7) in the presented results. Problem (7) can also be formulated in an unconstrained form by introducing a regularization parameter $\lambda$ which determines the weight of the $\ell_1$-norm penalty, as in a LASSO formulation $\min_{x} \|Hx - p\|_2^2 + \lambda \|x\|_1$.\(^{18}\)

A guarantee of the equivalence between the $\ell_0$ and $\ell_1$ minimization problems is that the underlying signal is sparse (a frequent scenario in array processing) and that the columns of the sensing matrix $H$ are mutually incoherent. A measure of column coherence is given by the largest non-diagonal element of the Gram matrix $(H^H H)^{15}$

$\mu(H) = \max_{i \neq j} |(H^H H)_{ij}|$. This relates to the restricted isometry property condition, $(1 - \delta_2)\|x\|_2^2 \leq \|Hx\|_2^2 \leq (1 + \delta_2)\|x\|_2^2$, and more specifically to the $\delta_2$ case, since $\delta_2 = \mu$. This aspect is not pursued further due to space limitations, but will be addressed in a full-length publication. The matrix coherence is an indicator of the sensitivity of the solution to noise and therefore robustness in a given scenario.

The estimated vector $\hat{x}$ that minimizes Eq. (7) comprises the amplitudes of the waves used in the expansion; based on their direction of incidence, it is possible to localize sound sources. If the aim were to reconstruct the sound-field elsewhere than measured, the obtained expansion can be used to extrapolate the sound field via a reconstruction matrix $H_s$ that relates the obtained expansion coefficients $\hat{x}$ to the desired reconstruction points in the medium $r_s$. See, for example, Refs. \(10\) and \(11\).

3. Experimental results

The proposed method is examined through an experiment in the large anechoic chamber at the Technical University of Denmark—DTU. An omnidirectional source is used (an Omnisource, Brüel&Kjær, Nærum, Denmark), placed 3.5 m away from the surface of a rigid spherical array of radius $a = 9.75$ cm (Fig. 1 left). The array consists of 64 B&K microphones, near-uniformly distributed over its surface, which can sample up to 7 orders of spherical harmonics.\(^{8}\) The source is driven with white noise, and a spectral resolution of 1 Hz is used for the analysis. The pressure on the array surface at 500 Hz ($ka \approx 0.9$) is shown in Fig. 1 (center). A plane wave expansion of 512 waves is considered [i.e., the plane-wave expansion described in Eq. (2)], and their angular direction is shown in Fig. 1 (right). The plane wave directions $(\Omega_{\theta,\phi})$ are distributed uniformly based on a Thomson problem, that considers equally charged particles on a sphere, therefore yielding a uniform sampling over the sphere. The aim of the method in this case is to use the plane wave expansion to detect the direction where the source is radiating from.

Figure 2 compares the conventional LS solution (left column), with the proposed CS solution (right column). The CS solution is obtained as in Eq. (7), and the LS solution too, but using the $\ell_2$-norm $\|x\|_2$ instead. The assumed relative upper bound of the noise floor is $\varepsilon = 5 \times 10^{-6}$ corresponding to a signal-to-noise ratio (SNR) of 25 dB. The top row

![Fig. 1](image-url)
of the figure shows the recovered pressures by each of the techniques. The center row shows the magnitude of the coefficients resulting from the two methodologies, showing that the set of coefficients obtained by the two approaches is significantly different. In the $l_2$ least-squares approach (Fig. 2, mid-left) all of the coefficients are non-zero, whereas the proposed CS method returns roughly two non-zero coefficients (Fig. 2, mid-right). The bottom row of the figure shows 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). It is noticeable that the LS solution corresponds to a conventional beamformer output, with its characteristic limited spatial resolution. The CS solution yields two coefficients, successfully detecting the direction of arrival of the wave radiated by the loudspeaker, with optimal accuracy. The computational cost of the CS solution is somewhat greater than an LS one (about 3 s compared to 0.3 s). Regarding the quantitative recovery, the power in each of the CS solution coefficients is larger than in the LS solution. As expected $\|\hat{x}_{l_2}\|_2 < \|\hat{x}_{l_1}\|_1$, although $\|\hat{x}_{l_1}\|_1 < \|\hat{x}_{l_2}\|_2$; This has significant consequences in the quantitative analysis of the sound field, both for estimating the power of the incident waves, and for reconstructing or reproducing the wave field captured by the array.

All in all, it is noticeable that the recovered sound pressure by the two methods is virtually identical, although the differences in the recovered coefficients are substantial. A few remarks about the proposed CS approach are outlined: (i) The spatial resolution is enhanced and approaches an ideal delta function; (ii) the sparsity
requirement acts as a “regularization” mechanism that suppresses noise; (iii) for a given noise level, the robustness of the CS solution can be anticipated from the incoherence of the columns of the sensing matrix, which is known a priori and thus unreliable estimates can be foreseen.

4. Conclusion

This study proposes and examines a method that makes use of the framework provided by CS for spherical array processing. The solution is obtained via $\ell_1$-minimization, which promotes sparsity on the solution, i.e., requires that only few coefficients are non-zero. Experimental results are presented. The results show that it is possible to identify the direction of arrival of the waves with remarkable spatial resolution and accuracy. The results indicate that it is possible to accurately represent the measured data with a minimum number of coefficients. This gives the method a promising perspective for different applications in acoustics such as direction of arrival estimation, sound field reconstruction, and sound field analysis.

Acknowledgments

This work was supported by the Danish Council for Independent Research (DFF), under Grant No. FTP/12-126364.

References and links