

ROOM HELPS: ACOUSTIC LOCALIZATION WITH FINITE ELEMENTS

Ivan Dokmanić and Martin Vetterli

School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
{ivan.dokmanic,martin.vetterli}@epfl.ch

ABSTRACT

Acoustic source localization often relies on the free-space/far-field model. Recent work exploiting spatio-temporal sparsity promises to go beyond these scenarios. However, it requires the knowledge of the transfer functions from each possible source location to each microphone. We propose a method for indoor acoustic source localization in which the physical modeling is implicit. By approximating the wave equation with the finite element method (FEM), we naturally get a sparse recovery formulation of the source localization. We demonstrate how exploiting the bandwidth leads to improved performance and surprising results, such as localization of multiple sources with one microphone, or hearing around corners. Numerical simulation results show the feasibility of such schemes.

Index Terms— Source localization, finite element method, sparse approximation, indoor localization, reverberant localization

1. INTRODUCTION

A common assumption in source localization methods such as beamforming, subspace methods or different parametric methods is that the sound propagates in free space [1]. Performance of these methods degrades substantially in the presence of multipath. Another frequent assumption is that the sources are in the far-field. This means that the wavefront generated by a source arrives at all microphones from the same direction.

These assumptions are essentially violated in rooms. We describe a method for localizing acoustic sources inside a room unaffected by this difficulty.

In fact, by correctly modeling the wave propagation effects, we can use the room to our advantage. In addition to estimating the source locations using a microphone array, we observe results such as localization of a pulse source using only one microphone, or localization of sources hidden behind corners.

We note that the idea of exploiting a known propagation model for source localization is known as matched-field processing, with early work in underwater acoustics already in the '70s [2].

1.1. Related Work

Two recent papers approach the source localization as a sparse reconstruction problem. In [3], the authors exploit the spatial sparsity of the sources. Since the sources are few, the vector of active locations is sparse and they reconstruct it using convex optimization. They also show how to treat wideband sources. Similarly, in [4] the authors assume the sources to be sparsely located on some grid.

This work was supported by an ERC Advanced Grant – Support for Frontier Research – SPARSAM Nr: 247006.

They further assume the temporal sparsity in a known dictionary and show how this assumption improves the estimation performance.

Both of these works assume the knowledge of transfer functions from all the possible source locations to all the possible microphone locations. In free space this modeling is easy, but in indoor spaces it could be a prohibitive requirement. Experiments therefore treat free-space/far-field situations, or ray simulations with a small number of reflections.

1.2. Main Contributions and Outline

We propose a method that directly relates acoustics with sparsity in arbitrary geometries. This is achieved by modeling the problem using the wave equation and then approximating the solution with the finite element method (FEM). We exploit the spatial sparsity of the sources, but assume no sparsity in the temporal domain. We assume however the knowledge of the room geometry.

In this paper we focus on the description of the physical aspect of the method. Our aim is to advocate what we consider to be a powerful concept. We are not concerned with the performance of associated sparse approximation methods and use them as an off-the-shelf technology with some modifications. We leave the discussion of further improvements, such as adaptive meshing and fine tuning of the sparse approximation algorithms, to a forthcoming extended version of this paper.

The paper is structured as follows. In Section 2 we describe the source model, the acoustic wave equation, and we derive the FEM matrix form. We formulate the source localization problem in the FEM context in Section 3, and furthermore show how to take advantage of the source bandwidth. We also show that knowing the source spectrum enables us to use simple linear inversion for localization. We verify the effectiveness of the method through numerical experiments in Section 4.

2. PROBLEM STATEMENT AND THE WAVE EQUATION

2.1. Setup

Consider K localized acoustic sources inside a room described by a region $D \subset \mathbb{R}^3$. The setup assumed throughout the paper is illustrated in Fig. 1.

Assume that the spatial distribution of the sources is given by a set points at locations $\{\mathbf{x}_k\}_{k=1}^K$, $\mathbf{x}_k \in D$. The k th source's waveform is given by a signal $s_k \in L^2_{\mathbb{R}}([0, \infty))$. These signals may represent music, speech or other arbitrary sounds. Total source distribution inside the room is then described by a function f ,

$$f(\mathbf{x}, t) = \sum_{k=1}^K s_k(t)\delta(\mathbf{x} - \mathbf{x}_k). \quad (1)$$

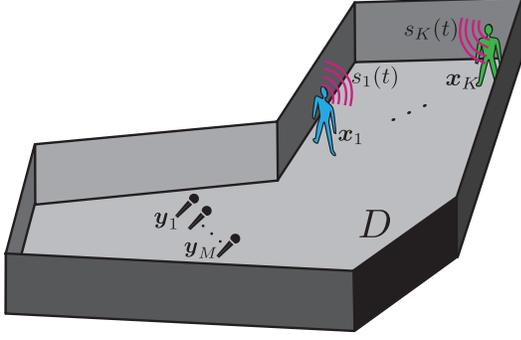


Fig. 1. Setup of the problem. We want to localize the acoustic sources emitting waveforms $\{s_k\}_{k=1}^K$, located at $\{\mathbf{x}_k\}_{k=1}^K$ with microphones at $\{y_m\}_{m=1}^M$ inside a known room D .

Sources generate pressure variations, which we denote by $u(\mathbf{x}, t)$. We observe $u(\mathbf{x}, t)$ with M microphones located at $\{y_m\}_{m=1}^M$ and attempt to solve the following problem.

Problem 1. Given access to measurements of sound pressure $\{u(y_m, t) + \epsilon_m(t)\}_{m=1}^M$ inside a known room D , where $\{\epsilon_m\}_{m=1}^M$ accounts for the modeling mismatch and noise, find the source locations $\{\mathbf{x}_k\}_{k=1}^K$.

2.2. Wave and Helmholtz Equation

Acoustic wave motion corresponds to changes of acoustic pressure around the mean value (often the atmospheric pressure) [5]. Under some fairly nonrestrictive assumptions, the pressure $u(\mathbf{x}, t)$ satisfies the following PDE, called the wave equation,

$$-\Delta u + \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = f, \quad (2)$$

with f being the source term.

Many applications do not require a full time-dependent wave equation. The system is analyzed under the assumption of a time-harmonic field $u(\mathbf{x}, t) = \hat{u}(\mathbf{x}, \omega)e^{-i\omega t}$, which is equivalent to taking the Fourier transform of the wave equation (2) with respect to time. This leads to the Helmholtz equation,

$$-\Delta \hat{u}(\mathbf{x}, \omega) - (\omega^2/c^2)\hat{u}(\mathbf{x}, \omega) = \hat{f}(\mathbf{x}, \omega). \quad (3)$$

Equation (3) does not involve time derivatives since the Fourier transform simplified them into a multiplication with the frequency squared.

Accounting for the source model (1), the Helmholtz equation (3) becomes

$$-\Delta \hat{u}(\mathbf{x}, \omega) - (\omega^2/c^2)\hat{u}(\mathbf{x}, \omega) = \sum_{k=1}^K \hat{s}(\omega)\delta(\mathbf{x} - \mathbf{x}_k), \quad (4)$$

where \hat{s} is the Fourier transform of s .

To have a complete characterization of the wave equation, we must specify the boundary conditions. In this paper we assume sound-hard walls corresponding to the Neumann boundary condition $\langle \nabla u(\mathbf{x}, t), \mathbf{n}(\mathbf{x}) \rangle = 0$, $\mathbf{x} \in \partial D$, where $\mathbf{n}(\mathbf{x})$ is the unit normal on the wall and $\langle \cdot, \cdot \rangle$ denotes the inner product. We note however that arbitrary impedance conditions are possible.

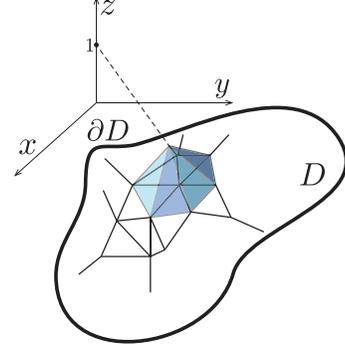


Fig. 2. Triangular mesh in a plane. Elements ϕ are pyramids of height 1.

2.3. Finite Elements for Helmholtz Equation

Consider now the Helmholtz equation (3). To arrive at the FEM formulation, we multiply both sides of the equation by a *test* function v and integrate over the room,

$$-\int_D \Delta \hat{u} v \, d\mathbf{x} - k^2 \int_D \hat{u} v \, d\mathbf{x} = \int_D \hat{f} v \, d\mathbf{x}. \quad (5)$$

Intuitively, if we require that this holds for all possible test functions v , then this form is equivalent to the original *pointwise* equation. Actually we require it to hold for all v that are *admissible*. For more details, see [6].

Equation (5) is asymmetric in that \hat{u} “has” second derivatives, while v “has” no derivatives. A more symmetric form is obtained after applying the Green’s theorem to the first integral,

$$\int_D \langle \nabla \hat{u}, \nabla v \rangle \, d\mathbf{x} - k^2 \int_D \hat{u} v \, d\mathbf{x} = \int_D \hat{f} v \, d\mathbf{x}. \quad (6)$$

Equation (6) is called the weak form of the Helmholtz equation. Note that the additional term produced by the application of Green’s theorem vanishes thanks to the boundary condition.

Let us now use the weak form to find an approximate solution $\hat{u}^* \approx \hat{u}$ as a linear combination of N *trial* functions $\{\phi_k\}_{k=1}^N$, $\hat{u}^*(\mathbf{x}, \omega) = \sum_{i=1}^N \hat{u}_{\omega,i}^* \phi_i(\mathbf{x})$. Plugging \hat{u}^* into the weak form (6), we get one linear equation in N unknowns $\{\hat{u}_{\omega,i}^*\}_{i=1}^N$. The problem is now reduced to the computation of these coefficients. To obtain the necessary N equations, we simply pick N *test* functions v_1, \dots, v_N .

Putting the pieces together in (6), the weak form becomes

$$\begin{aligned} \int_D \left\langle \nabla \left(\sum_{i=1}^N \hat{u}_{\omega,i}^* \phi_i \right), \nabla v_j \right\rangle \, d\mathbf{x} - \frac{\omega^2}{c^2} \int_D \left(\sum_{i=1}^N \hat{u}_{\omega,i}^* \phi_i \right) v_j \, d\mathbf{x} \\ = \int_D \hat{f} v_j \, d\mathbf{x}, \quad 1 \leq j \leq N. \end{aligned} \quad (7)$$

For each j we have a linear equation with the unknowns $\hat{u}_{\omega,1}^*, \dots, \hat{u}_{\omega,N}^*$. Written more compactly, the system is

$$\sum_{i=1}^N [K_{i,j} - (\omega^2/c^2)M_{i,j}] \hat{u}_{\omega,i}^* = \hat{f}_j, \quad 1 \leq j \leq N. \quad (8)$$

where $K_{i,j} = \int_D \langle \nabla \phi_j, \nabla v_i \rangle \, d\mathbf{x}$, $M_{i,j} = \int_D \phi_i v_j \, d\mathbf{x}$ and $\hat{f}_j =$

$\int_D \widehat{f} v_i \, d\mathbf{x}$, or in a matrix form,

$$[\mathbf{K} - (\omega^2/c^2)\mathbf{M}]\widehat{\mathbf{u}}_\omega^* = \widehat{\mathbf{f}}_\omega. \quad (9)$$

A common choice is $v_j = \phi_j$ so that \mathbf{K} and \mathbf{M} are symmetric positive-definite matrices.

Interestingly, the approximate solution $\widehat{\mathbf{u}}^*$ is an orthogonal projection of the exact solution $\widehat{\mathbf{u}}$ onto the linear subspace spanned by ϕ_1, \dots, ϕ_N . This projection is called Galerkin's projection in honor of the Russian mathematician Boris Galerkin. From Galerkin's idea to FEM there is only one small step—one chooses the trial functions $\{\phi_i\}_{i=1}^N$ to be localized piecewise polynomials.

There are two immediate benefits from having this formulation. First, if we choose ϕ 's to have a localized support, many of the elements will not overlap. This means that many integrals for $K_{i,j}$ and $M_{i,j}$ will be zero, and $\mathbf{K} - (\omega^2/c^2)\mathbf{M}$ will be sparse. Second, since there are no more second derivatives, we can safely choose piecewise linear elements. Typically, we discretize the domain with a triangular mesh, and use piecewise linear elements centered at mesh nodes. A 2-D illustration is given in Fig. 2.

3. SOURCE LOCALIZATION WITH FEM

The salient point of having finite elements for test and trial functions is their restricted spatial support. Given a sufficiently fine mesh, if we measure the amplitude of the pressure oscillations at some location \mathbf{x} at a frequency ω , we approximately measure the value of the coefficient $\widehat{u}_{\omega,i}^*$ corresponding to the finite element centered around \mathbf{x} .

Let $\mathbf{A}_\omega \stackrel{\text{def}}{=} \mathbf{K} - (\omega^2/c^2)\mathbf{M}$ be the matrix of a FEM-discretized Helmholtz equation. For convenience, let also $\mathbf{G}_\omega \stackrel{\text{def}}{=} \mathbf{A}_\omega^{-1}$. Then given the source distribution $\widehat{\mathbf{f}}$ expanded in the chosen FE basis, the solution $\widehat{\mathbf{u}}_\omega^*$ is obtained as $\widehat{\mathbf{u}}_\omega^* = \mathbf{G}_\omega \widehat{\mathbf{f}}_\omega$. Let us assume that the microphones are located at the mesh nodes. This is not unrealistic, since we know the array geometry by design, and we can always mesh in such a way that this is true. Denoting the set of indices corresponding to $\{\mathbf{y}_m\}_{m=1}^M$ by \mathcal{Y} , we have

$$\widehat{\mathbf{u}}_\omega^*[\mathcal{Y}] = \mathbf{G}_\omega[\mathcal{Y}, :] \widehat{\mathbf{f}}_\omega, \quad (10)$$

where the indexing with \mathcal{Y} selects the rows with indices in \mathcal{Y} . Since the sources are spatially sparse, only a small fraction of elements in $\widehat{\mathbf{f}}$ are non-zero—a consequence of the localization of finite elements.

3.1. Point Source Emitting a Sinusoid

Before explaining the general case we examine the simpler case of a single point source emitting a single tone. Source localization is now reduced to finding a single non-zero element in $\widehat{\mathbf{f}}_\omega$ if the source is at the mesh node, and a small cluster of nonzeros if it is not.

A simple solution is to find in \mathbf{G}_ω the column proportional to $\widehat{\mathbf{u}}_\omega^*$. Writing $\mathbf{G}_\omega = [\mathbf{g}_{\omega,1}, \dots, \mathbf{g}_{\omega,N}]$, the solution is obtained as

$$\begin{aligned} \widehat{\mathbf{f}} &= \alpha \mathbf{e}_i, \quad \text{where} \\ i &= \arg \min_j \|\widehat{\mathbf{u}}^*[\mathcal{Y}] - \frac{1}{\|\mathbf{g}_j\|^2} \langle \widehat{\mathbf{u}}^*[\mathcal{Y}], \mathbf{g}_j \rangle \mathbf{g}_j\|, \text{ and} \\ \alpha &= \langle \widehat{\mathbf{u}}^*[\mathcal{Y}], \frac{1}{\|\mathbf{g}_i\|} \mathbf{g}_i \rangle, \end{aligned} \quad (11)$$

and \mathbf{e}_i is the i th canonical basis vector in \mathbb{R}^N .

If the noise is a zero-mean Gaussian noise, (11) is the maximum likelihood solution, given that the source is at the mesh node.

3.2. Multiple Wideband Point Sources

In the presence of multiple sources, we are no longer trying to find the best column, but rather the best selection of columns that explains our measurements.

Additionally, for wideband sources, we have an important observation: the Helmholtz equation with the assumed source model (4) is valid for all frequencies ω . We can record $u(\mathbf{x}, t)$, and then compute its Fourier transform to get $\widehat{u}(\mathbf{x}, \omega)$. We thus obtain the solution of the Helmholtz equation for many ω 's.

For each of these frequencies we know the matrix \mathbf{G}_ω and we have measured some entries in $\widehat{\mathbf{u}}_\omega^*$. If we choose a discrete set of frequencies $\{\omega_i\}_{i=1}^F$, we arrive at the following system of matrix equations,

$$[\mathbf{K} - (\omega_i^2/c^2)\mathbf{M}]\widehat{\mathbf{u}}_\omega^* = \widehat{\mathbf{f}}_i, \quad 1 \leq i \leq F, \quad (12)$$

where we only know (measure) a couple of entries in each $\widehat{\mathbf{u}}_{\omega_i}^*$. In the established notation, this means that we are solving the following system,

$$\begin{cases} \widehat{\mathbf{u}}_{\omega_1}^*[\mathcal{Y}] &= \mathbf{G}_{\omega_1}[\mathcal{Y}] \widehat{\mathbf{f}}_{\omega_1} \\ \vdots \\ \widehat{\mathbf{u}}_{\omega_F}^*[\mathcal{Y}] &= \mathbf{G}_{\omega_F}[\mathcal{Y}] \widehat{\mathbf{f}}_{\omega_F}, \end{cases}$$

but, importantly, in each of the subproblems we are searching for a sparse solution with the identical sparsity pattern. This is because we assume that the sources do not move over the observation period.

This problem is in a way complementary to already well investigated sparse recovery algorithms, and many of them can be mechanically extended to our scenario. Studying the performance of these algorithms is outside of the focus of this paper. For our experiments section we used a naive extension of the orthogonal matching pursuit [7], where we compute the *goodness* of a column by summing its goodnesses over all F subproblems. Even if it is suboptimal, this simple scheme yields interesting results.

3.3. Localizing Pulse Sources with One Microphone

Consider the case where all the waveforms $s_k(t)$ are equal, known and wideband (say the sources emit a pulse). Then something curious happens. Let there only be one microphone located at \mathbf{y} and let the corresponding element index be n . For a frequency ω we have

$$\widehat{\mathbf{u}}_\omega^*[n] = \mathbf{G}_\omega[n, :] \widehat{\mathbf{f}}, \quad (13)$$

but now $\widehat{\mathbf{u}}_{\omega_1}[n]$ is just a scalar, and $\mathbf{G}_{\omega_1}[n, :]$ is a row vector. Assume that $s_k(t) = \delta(t)$. Since $\int_{\mathbb{R}} \delta(t) e^{-i\omega t} dt = 1$, we know that the source vector $\widehat{\mathbf{f}}_\omega = \widehat{\mathbf{f}}$ remains constant for all frequencies. Now repeat this for many frequencies. This means that we can stack the observations and the matrix rows to obtain the following equation,

$$\begin{bmatrix} \widehat{\mathbf{u}}_{\omega_1}^*[n] \\ \vdots \\ \widehat{\mathbf{u}}_{\omega_F}^*[n] \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{\omega_1}[n, :] \\ \vdots \\ \mathbf{G}_{\omega_F}[n, :] \end{bmatrix} \widehat{\mathbf{f}}. \quad (14)$$

As soon as the rank of the matrix on the right hand side is at least N we recover $\widehat{\mathbf{f}}$ by an ordinary linear inversion. So, at least in theory, we can recover an arbitrary number of pulse sources using one microphone and solving a linear system.

Without discussing obvious practical issues with performing this computation, such as conditioning, we think that the concept is quite interesting in its own right.

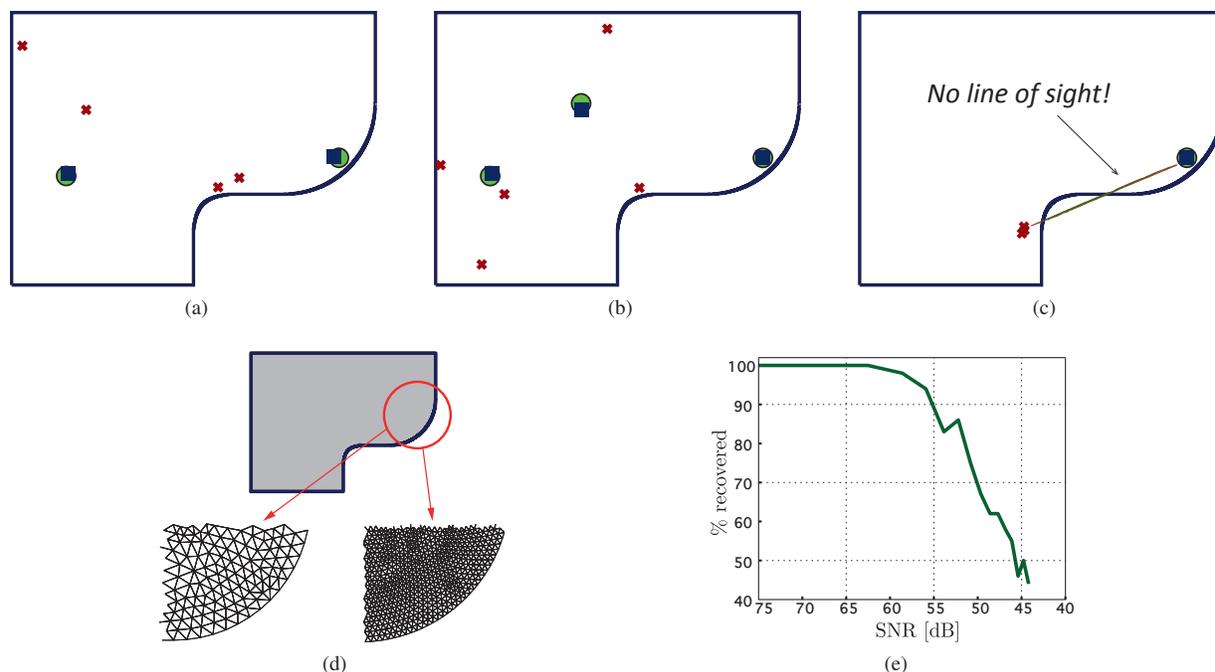


Fig. 3. Localizing K sources with M microphones: red crosses represent microphones, green circles are true source locations, blue squares are estimated locations; (a) $K = 2$, $M = 4$; (b) $K = 3$, $M = 5$; (c) $K = 1$, $M = 3$, *no line of sight*; (d) coarse (estimation) and fine (simulation) mesh used in experiments; (e) performance of the localization in noise.

4. NUMERICAL EXPERIMENTS

We have validated the theoretical results on a number of numerical simulations. To ensure a realistic validation, we use different meshes for the simulation of the field and for the source localization. For visualization we use a 2-D room, but the developed theory works both in 2-D and in 3-D.

Acoustics are simulated using a very fine mesh, and this result is considered to be the true acoustic wavefield. The estimation algorithm is run on a different, considerably coarser mesh. The two meshes are illustrated in Fig. 3d). As a matter of fact, the estimation mesh is coarser than what we would use in a real situation.

Fig. 3a) and Fig. 3b) show the localization of 2 and 3 wideband sources with 4 and 5 microphones. We observe that even with the relatively large model mismatch due to the coarse mesh, the sources are correctly located. In Fig. 3c) we show the localization of one source with three microphones. Unlike what you might expect from classical methods, the source is successfully located even if there is no direct channel between it and the array. The room helps us to find the source.

We have also experimented with adding noise to measurements. Rate of success in localizing one source with an array of 5 microphones is shown in Fig. 3e) for different noise levels.

5. CONCLUSION

We have proposed a method for acoustic source localization based on the acoustic wave equation and the finite element method (FEM). The method exploits the implicit physical modeling provided by FEM, and the source sparsity through the sparse approximation methods. We have also shown how the source bandwidth may be

used to better condition the problem and reduce the number of microphones needed. If the sources emit a pulse, we have a possibility of estimating their locations using linear inversion. Numerical experiments confirm the effectiveness of the proposed methods, but additional research is needed to improve the performance in noise and to assess the sensitivity to imperfect geometry knowledge. Ongoing work includes adaptive meshing for increased resolution and developing sparse recovery algorithms specifically for this scenario.

6. REFERENCES

- [1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, July 1996.
- [2] H. P. Bucker, "Use of calculated sound fields and matched-field detection to locate sound sources in shallow water," *J. Acoust. Soc. Am.*, vol. 59, no. 2, pp. 368–373, 1976.
- [3] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.
- [4] D. Model and M. Zibulevsky, "Signal reconstruction in sensor arrays using sparse representations," *Signal Process.*, vol. 86, no. 3, pp. 624–638, Mar. 2006.
- [5] P. M. Morse and U. K. Ingard, *Theoretical Acoustics*, Princeton University Press, New Jersey, 1968.
- [6] G. Strang and G. J. Fix, *An Analysis of the Finite Element Method*, Prentice-Hall series in automatic computation. Prentice-Hall, New Jersey, 1973.
- [7] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proc. Ann. Asilomar Conf. Signals, Syst., Comput.*, Nov. 1993, pp. 40–44.