Near-field acoustic imaging based on Laplacian sparsity

Efren Fernandez-Grande\(^{(a)}\) and Laurent Daudet\(^{(b)}\)

\(^{(a)}\)Acoustic Technology, Department of Electrical Engineering, Technical University of Denmark, efg@elektro.dtu.dk
\(^{(b)}\)Institut Langevin, Paris Diderot University, ESPCI and CNRS UMR 7587, Paris, France, laurent.daudet@espci.fr

Abstract:

We present a sound source identification method for near-field acoustic imaging of extended sources. The methodology is based on a wave superposition method (or equivalent source method) that promotes solutions with sparse higher order spatial derivatives. Instead of promoting direct sparsity, as in standard compressive sensing or basis pursuit approaches, solutions with a piecewise constant gradient or curvature are promoted, suitable for modeling extended sources that are subject to smooth spatial variations. The obtained results are compared to Least Squares and Compressive Sensing solutions, and the validity of the wave extrapolation used for the reconstruction is examined. It is shown that this methodology can overcome conventional limits of spatial sampling, and is therefore valid for wide-band acoustic imaging of extended sources.

Keywords: Sound radiation, acoustic array processing, near-field acoustic holography, compressive sensing
Near-field acoustic imaging based on Laplacian sparsity

1 Introduction

Near-field acoustic holography and near-field reconstruction methods are useful to examine how acoustic sources radiate sound into a medium [1, 2]. These methods rely on sampling the wavefield radiated by a source, and using a wave model (either based in k-space [1, 2], plane waves [3], point sources [4], spherical harmonics [5, 6, 7], etc.) to represent the sound field and infer all acoustic quantities in the volume near the source: pressure, velocity, intensity, and sound power.

One of the main limitations of acoustic reconstruction methods and acoustic imaging techniques is the need to have a sufficiently dense spatial sampling, in order to prevent aliasing errors. For this reason, there has been increasing interest in methods based on Compressive Sensing, Compressed Sensing or Compressive Sampling (CS) [8, 9], as the methodology makes it possible to overcome the conventional Nyquist sampling limit. This framework has been successfully applied to localization methods since nearly a decade [10, 11, 12, 13]. It has been recently applied to acoustic holography and near-field reconstruction problems, first by Chardon et al. (2012) [14], using a conventional explicit discrete Fourier transform, later by Hald (20014) [15], and Fernandez-Grande and Xenaki (2015) [?] using the method of wave superposition or Equivalent Source Method (ESM) [16, 17, 4].

However, one fundamental assumption of CS is that the underlying problem is sparse, and this is generally not the case when examining the acoustic near-field of wave-bearing structures or spatially extended sources, as the regions that radiate sound may cover a large solid angle, thus violating the sparsity condition. This can result in solutions that are prone to high errors due to the underlying model mismatch [?]. It is the motivation of this paper to propose an alternative methodology, based on a total generalized variation framework [18, 19], that can overcome the sampling limit, and which is physically meaningful for the purpose of near-field imaging or imaging of spatially extended sources.

In this study we examine the possibility of representing the radiation of spatially extended sources by solutions with sparse spatial derivatives, rather than solutions that are directly sparse in space. Total Variation (TV) [20, 21] is a well-known approach that promotes solutions with a sparse gradient (i.e. few kinks in the solution). This results in a piece-wise constant representation of the source field, i.e. regions of equal amplitude, devoid of rapid noisy fluctuations and without ringing artifacts near the edges of the source. Nonetheless, it seems more meaningful to consider the second order spatial derivative, based on a Total General Variation framework [18, 19], since this will result in a linear piece-wise representation. This representation is more meaningful in the general case of acoustic sources, and will preserve a high spatial resolution (or spatial definition), avoiding staircasing effects as in conventional TV solutions. To our best knowledge this approach has not previously been used in acoustics before.

In this paper we consider a fused approach, where spatial resolution is further enhanced by
combining it with direct space-domain sparsity via an additional $\ell_1$-norm penalty. We present the theory of the method and a numerical study to assess its validity.

2 Theory

The method of wave superposition, or equivalent source method (ESM), consists on representing the sound field radiated by an acoustic source as the superposition of the sound fields radiated by a combination of point sources

$$p(r,\omega) = j\omega \rho \int_S U(r_0) G(r,r_0) dS(r_0),$$

(1)

where $p(r,\omega)$ is the sound pressure at an observation point $r$ and angular frequency $\omega$ (the explicit time dependency $e^{j\omega t}$ is omitted). The function $U(r_0)$ is the surface velocity on the equivalent source surface $S(r_0)$, and $G(r,r_0)$ is the free-field Green's function between the equivalent sources at $r_0$ and the observation point $r$ (note that dipoles and other multipoles could also be considered in this framework). The particle velocity vector $u(r)$, from Euler's equation of motion, is

$$u(r) = -\int_S q(r_0) \nabla G(r,r_0) dS(r_0).$$

(2)

It follows from Eqs. (1) and (2) that any arbitrary sound field can be fully represented by such equivalent source model. If a sound field is sampled in space and time via measurements with an array of microphones, the sound pressure at the microphone positions can be expressed as due to the radiation from the 'equivalent sources' distributed on the surface of the actual source (see Fig. 1; in practice, the equivalent sources are retracted from the surface of the source to prevent the singularity). Note that when the equivalent source model is discretised, the surface velocity $U(r_0)$ in (1) and (2) results into the volume velocities $Q_n$ of each of the $n$ equivalent sources. Equation (1) is discretised as

$$p = Gq + n$$

(3)

where $p = [p(r_1), p(r_2), ..., p(r_M)]^T \in \mathbb{C}^M$ is the measured pressure at the microphone positions, $q = [q(r_{0,1}), q(r_{0,2}), ..., q(r_{0,N})]^T \in \mathbb{C}^N$ is the coefficient vector containing the strength of the sources, which relate to their volume velocity as $q_i = j\omega \rho Q_i$, and $G \in \mathbb{C}^{M \times N}$ contains the entries from the free-field Green's function between the position of the microphones $r_m$ and equivalent sources $r_{0,n}$. We consider additive noise $n$ with variance $\varepsilon$. It should be noted that this problem is typically underdetermined $(N >> M)$.

In order to estimate the unknown coefficient vector $q$, the system of equations can be formulated as a constrained optimization problem of the form

$$\hat{q} = \arg \min_q ||q||_p \quad \text{subject to} \quad ||Gq - p||_2 \leq \varepsilon,$$

(4)

where $|| \cdot ||_p$ is the $\ell_p$-norm defined as $||x||_p = \left(\sum |x|^p\right)^{1/p}$. This problem is usually solved in a least squares sense, which consists of choosing the $\ell_2$-norm in (4), yielding the solution with minimum energy from the solution subspace.
The Compressive Sensing (CS) methodology consists on solving (4) via minimizing the $\ell_1$-norm [8, 9]. CS provides the sparsest solution to the problem, provided that the columns of the transfer matrix are mutually incoherent, and that the problem is in fact inherently sparse (i.e. the coefficient vector $q$ has few non-zero elements). One of the remarkable benefits of CS is that it can overcome the Nyquist sampling limit, as the sparse solution is devoid of sidelobes, and therefore aliasing effects. However, this is problematic when addressing continuous sources, as these are not sparse [?].

In this study we propose to obtain a solution $\hat{q}$ based on sparse gradients or higher order derivatives instead of direct spatial sparsity. Let us consider the generalised formulation of (4),

$$\hat{q} = \arg \min_q ||Dq||_1 \quad \text{subject to } ||Gq - p||_2 \leq \epsilon,$$

(5)

where $D$ is a weighting matrix. The so-called Total Variation (TV) problem consists on a gradient finite-difference operator $Dq = \nabla q$ [20] penalizing the number of non-zero gradients in the coefficient solution. This promotes piecewise constant solutions in space, suitable for ‘flat’ or constant amplitudes, such as pistons. However, the approach is ill-suited for representing smoothly varying deflection shapes, as commonly found in wave-bearing structures. Therefore, it is here proposed to use a Laplacian penalty

$$\hat{q} = \arg \min_q ||\nabla^2 q||_1 \quad \text{subject to } ||Gq - p||_2 \leq \epsilon,$$

(6)

which will promote a piecewise linear solution. In this study a fused approach is used, in which in addition to promoting Laplacian sparsity [as in Eq. (6)], direct space-domain sparsity is promoted via direct $\ell_1$-norm minimization of the solution vector. The corresponding weighting matrix in (5) is

$$D = \begin{bmatrix} D_2 \\ I \end{bmatrix},$$

(7)

with $D \in \mathbb{C}^{2N \times N}$, $I$ is the identity matrix, and $D_2$ is a Laplacian finite-difference operator of second order (in this study implemented on a 9-point stencil). Problem (5) with the weighting matrix $D$ in Eq. (7) is used in the presented results. This approach results in sensible solutions for spatially extended acoustic sources, as well as localized points in space.
3 A simulation study

3.1 Sound field reconstruction

A numerical simulation is used to examine the proposed methodology. The solution from the proposed TGV method, as in Eqs. (5) and (7) is examined. For comparison, also the Least-Squares solution (LS) and Compressed Sensing solution (CS) are considered. First, a vibrating steel plate is considered, of dimensions $50 \times 50$ cm$^2$, and thickness 2 mm. The plate is centred in the origin of coordinates, and driven by a point force at 600 Hz. The array used for the simulated measurement is a 60 channel array, with pseudo random sampling in space of diameter 50 cm. The measurement takes place 8 cm away from the plate, and the reconstruction 5 cm away from it. An equivalent source mesh of $30 \times 30$ sources is used, distributed over a $50 \times 50$ cm$^2$ rectangular grid. The equivalent sources are positioned in the plane $z = 0$ (i.e. on the plate’s surface). Additive white noise is included in the simulated measurements with a signal-to-noise ratio (SNR) of 30 dB. The noise floor estimate $\varepsilon$ is selected accordingly.

![Figure 2: Estimated coefficients for a vibrating plate; 600 Hz, $z_h = 8$ cm, $z_s = 5$ cm, $z_0 = 0$](image)

Figure 2 shows the magnitude of the estimated coefficients for the three methodologies (CS, LS and TGV). It is apparent how the CS reconstruction does not recover a spatially extended source, but rather a combination of point sources whose radiation approximates the observed data. The conventional LS method provides a spatially extended source distribution, indicating the spatial extent of the vibrating plate. The solution from the proposed TGV methodology is similar to LS, in the sense that a spatially extended source is recovered, although the spatial features are somewhat better resolved than LS, particularly towards the corners. The error on the reconstructed sound pressure field in $z = 5$ cm is of 9.9 % for the CS, 8.3 % for the LS, and 7.2 % for the TGV based reconstructions. When analysing the reconstructed normal velocity fields, the spatially averaged reconstruction error is 32 % for the CS based reconstruction, 18 % for the LS based reconstruction, and 16 % for the TV based reconstruction. The TGV and LS methods provide accurate reconstructions, since they recover an extended source. Contrarily, the CS approach has a greater error, because the problem addressed is not sparse, and the recovered solution suffers from an inherent model mismatch (i.e. the recovered coefficients are not a good model for representing the sound field radiated by the plate). This is particularly critical in the reconstructed velocity field. All in all, the TGV methodology is consistently the most accurate.
We consider the case of a monopole in order to analyse the response from the different methodologies (in the form of point-spread-functions). A monopole placed at \((x,y,z) = (-4,0,0)\) cm with volume velocity \(Q = 10^{-5} \text{ m}^3\text{s}^{-1}\) is considered. The same simulation setup as with the plate is used. Figure 3 shows the estimated solutions, depicting the magnitude of the coefficients through a vertical line \((-0.05,y)\) passing near the location of the point source. As expected, the CS solution recovers the point source almost perfectly, due to the inherent sparsity of the wave field. On the other hand, the LS solution results in a characteristic array response consisting of a mainlobe and sidelobes (dependent on the wavelength, aperture size, array spatial sampling and source distance). The TGV solution recovers a source with a certain spatial extent, with a finer spatial resolution than LS, but more importantly without sidelobes in its response. When considering the reconstructed fields, the spatially averaged error over the aperture of the reconstructed sound pressure is of CS: 5.7 \%, LS: 12 \%, and TGV: 10 \%. The error of the reconstructed normal velocity fields in this case are of CS: 7 \%, LS: 50 \%, and TGV: 32 \%, demonstrating the aforementioned differences.

3.2 Frequency and aliasing limit

We examine the frequency dependence of the proposed TGV method, and continue using LS and CS as a benchmark. As discussed in Sect. I, the influence of aliasing errors is critical for reconstruction methods. A longitudinal quadrupole source is considered for this study (we did not use a vibrating plate due to the discretization errors at high frequency). The quadrupole source is placed 8 cm away from the array, and the reconstruction is 5 cm away from the source. The equivalent sources, \(41 \times 41\), uniformly distributed over an aperture of \(50 \times 50 \text{ cm}^2\), are placed on the plane of the quadrupole. Additive white noise with SNR of 30 dB is added.
to the simulated measurements.

Figure 4: Spatially averaged error (mean over 10 realizations) for each frequency of the reconstructed pressure and normal velocity; \( z_h = 8 \) cm, \( z_s = 5 \) cm, \( z_0 = 0 \).

Figure 4 shows the results of the spatially averaged reconstruction error for the pressure and particle velocity fields with the three methods, from 100 Hz to 5 kHz, every 100 Hz. At each frequency there are 10 realizations of the reconstruction. The figure shows the mean (the standard deviation of the three methods is similar, somewhat greater for CS, but of the order of the background noise in any case). It is clear from Fig. 4 that the LS method breaks down at higher frequencies due to high sidelobe contamination, stemming from the spatial sampling limits. The CS provides an accurate wideband reconstruction as expected (the wideband validity of CS has been extensively exploited [8, 9]), and with a low error, because the actual source under study consists of point sources. The proposed TGV methodology is accurate at high frequencies, beyond the conventional sampling limits (as is the case also for CS), indicating its validity for wideband imaging.

Figure 5: Longitudinal quadrupole radiating at 5000 Hz. Estimated coefficients \( \hat{q} \) based on Compressive Sensing (left), Least Squares (center) and Total Generalized Variation (right). The quadrupole is located in the plane \( z_0 = 0, z_h = 8 \) cm.

The recovered solutions (CS, LS and TGV) for the quadrupole radiating at 5 kHz are shown in Fig. 5. It is apparent that the LS solution is contaminated by sidelobes of high level, which
explains the poor reconstruction at high frequencies. Contrarily the CS and the proposed TGV methodology can successfully reconstruct the sound field successfully over the entire frequency range, from 100 to 5000 Hz (with the same 60 microphones array). It is in fact interesting to note that the error for the CS and TGV methods decreases with frequency, essentially because the near-field from the quadrupole is effectively less dominant at higher frequencies, and the observed data is less redundant.

4 Conclusions
This study presents a method based on Total Generalized Variation (TGV), which is suitable for imaging spatially-extended acoustic sources. The method seeks to model an acoustic source as a superposition of elementary sources with sparse spatial derivatives. The framework of the method has been presented and the case of a ‘fused’ second order derivative has been examined numerically. The results indicate that the TGV method can provide broadband imaging, unlike the classical least-squares, and it is not restricted to spatially sparse problems, as the compressive sensing method is. These findings show good potential for acoustic and vibration measurements on spatially extended sources.

Acknowledgements
This work was supported by the Danish Council for Independent Research (DFF), under the individual postdoctoral grant FTP/12-126364. The first author would like to thank Niccolo Antonello and Angeliki Xenaki for discussions and input at the early stages of this work.

References


