



Compressive acoustic holography with block-sparse regularization

Fernandez Grande, Efren; Daudet, Laurent

Published in:
Journal of the Acoustical Society of America

Link to article, DOI:
[10.1121/1.5042412](https://doi.org/10.1121/1.5042412)

Publication date:
2018

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

[Link back to DTU Orbit](#)

Citation (APA):
Fernandez Grande, E., & Daudet, L. (2018). Compressive acoustic holography with block-sparse regularization. *Journal of the Acoustical Society of America*, 143(6), 3737-3746. <https://doi.org/10.1121/1.5042412>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Compressive acoustic holography with block-sparse regularization

Efren Fernandez-Grande, and Laurent Daudet

Citation: *The Journal of the Acoustical Society of America* **143**, 3737 (2018); doi: 10.1121/1.5042412

View online: <https://doi.org/10.1121/1.5042412>

View Table of Contents: <http://asa.scitation.org/toc/jas/143/6>

Published by the *Acoustical Society of America*

Compressive acoustic holography with block-sparse regularization

Efren Fernandez-Grande^{1,a)} and Laurent Daudet²

¹*Department of Electrical Engineering, Acoustic Technology, Technical University of Denmark (DTU), Building 352, DK-2800 Kgs. Lyngby, Denmark*

²*Institut Langevin, Paris Diderot University, ESPCI and CNRS, UMR 7587, Paris, France*

(Received 9 December 2017; revised 23 May 2018; accepted 25 May 2018; published online 29 June 2018)

Sparse reconstruction methods, such as Compressive Sensing, are powerful methods in acoustic array processing, as they make wideband reconstruction possible. However, when addressing sound fields that are not necessarily sparse (e.g., in acoustic near-fields, reflective environments, extended sources, etc.), the methods can lead to a poor reconstruction of the sound field. This study examines the use of sparse analysis priors to promote block-sparse solutions. In particular, a Fused Total Generalized Variation (F-TGV) method is developed, to analyze the sound field in the near-field of acoustic sources. The method promotes sparsity both on the spatial derivatives of the solution and on the solution itself, thus seeking solutions where the non-zero coefficients are grouped together. The performance of the method is examined numerically and experimentally, and compared with established methods. The results indicate that the F-TGV method is suitable to examine both compact and spatially extended sources. The method is promising for its generality, robustness to noise, and the capability to provide a wideband reconstruction of sound fields that are not necessarily sparse. © 2018 Acoustical Society of America. <https://doi.org/10.1121/1.5042412>

[PG]

Pages: 3737–3746

I. INTRODUCTION

Microphone array processing is a powerful means to address acoustic problems related to the spatial properties of sound. Array methods are useful to localize sound sources, infer properties of the sound field in enclosures, or analyze the sound field radiated by acoustic sources. In recent years, there has been a growing interest in microphone array methods based on Compressive Sensing (CS),^{1–4} as they make it possible to extend the frequency range of validity of the methods. Consequently, the spatial sampling requirements are alleviated, leading to making wideband reconstruction possible. CS has been employed successfully in source localization,^{5–7,23} sound field reproduction,^{8,9} near-field reconstruction problems,^{10–12} and characterization of material properties.¹³

CS exploits the underlying sparse structure of a problem to seek a perfect signal reconstruction.² However, many of the problems encountered in acoustics are not spatially sparse (e.g., in rooms, in the near-field of wave-bearing structures, or spatially extended sources). In near-field acoustic holography (NAH),^{14–17} which deals with the identification and reconstruction of the acoustic field near a source, the lack of source sparsity is typically inevitable. This has led to applying CS in transformed domains, such as the wavenumber domain¹⁰ and eigen-function domain.¹⁸ Recent studies^{11,12} show that CS can be successfully used to reconstruct sound fields in the near-field of a source, even when this source is not spatially sparse.¹² However, it leads to a source representation that is not physically significant, resulting in reconstruction errors.

In this work we aim at developing methods that promote block-sparsity instead of direct sparsity in space domain. The objective is to develop a method that preserves the favorable properties of sparse processing methods (most notably their capability to achieve a wideband reconstruction), and that is suitable for characterizing sources that are spatially extended, or not necessarily sparse. Taking as a starting point the Total Variation (TV) regularization method,^{19,20} we propose a Total Generalized Variation (TGV) framework to reconstruct the sound field near a vibrating source.^{21,22}

TV methods exploit the use of spatial gradients on the solution, to promote piece-wise constant estimates. These methods have been widely used for image reconstruction and de-noising,^{19,24–26} and recently in acoustics.²⁷ In this study we exploit the higher-order spatial derivatives of the solution estimate, as in TGV methods.^{21–23} In particular, the use of a Laplacian operator is examined, which promotes a piece-wise linear solution instead of a piece-wise constant one, as in conventional TV. We propose a Fused Total Generalized Variation (F-TGV) method, which combines direct spatial sparsity with a second order TGV (second derivative on the solution). The proposed method promotes block-sparse solutions: the solution estimate is not necessarily strictly sparse, but all the non-zero coefficients are grouped together. The F-TGV method can be understood as an analysis sparsity prior, as opposed to a synthesis one, because a specific structure is not imposed on the signal directly, but rather on a transformed domain. To the author's knowledge, it is the first time that such a method is proposed.

The paper is structured as follows: in Sec. II the methodology is presented, along with existing reconstruction methods. In Sec. III a numerical study is conducted that examines

^{a)}Electronic mail: efg@elektro.dtu.dk

the accuracy, robustness, and wideband properties of the methods. The sources used are a vibrating piston, a longitudinal quadrupole, and a monopole source. Finally in Sec. IV, an experimental study is presented that examines the radiation of a dipolar source, and examines the quantitative accuracy of the method via an estimation of the radiated sound power.

II. THEORY

A. ESM

The equivalent source method (ESM) is a common method used in NAH, sound radiation, and scattering problems.^{17,28–30} The fundamental idea is that an arbitrary sound field radiated by a sound source can be expressed as the superposition of the sound fields radiated by a combination of elementary point sources,

$$p(\mathbf{r}, \omega) = j\omega\rho \int_S U_n(\mathbf{r}_0)G(\mathbf{r}, \mathbf{r}_0)dS(\mathbf{r}_0), \quad (1)$$

where $p(\mathbf{r}, \omega)$ is the sound pressure at an observation point \mathbf{r} and angular frequency ω (the explicit time dependency $e^{j\omega t}$ is omitted). The function $U_n(\mathbf{r}_0)$ is the apparent velocity on the equivalent source surface $S(\mathbf{r}_0)$, and $G(\mathbf{r}, \mathbf{r}_0)$ is the free-field Green's function between an equivalent source at \mathbf{r}_0 and the observation point \mathbf{r} ,

$$G(\mathbf{r}, \mathbf{r}_0) = \frac{e^{-jk(|\mathbf{r}-\mathbf{r}_0|)}}{4\pi|\mathbf{r}-\mathbf{r}_0|}. \quad (2)$$

The particle velocity vector $\mathbf{u}(\mathbf{r}, \omega)$ is

$$\mathbf{u}(\mathbf{r}, \omega) = - \int_S U_n(\mathbf{r}_0)\nabla G(\mathbf{r}, \mathbf{r}_0)dS(\mathbf{r}_0). \quad (3)$$

This framework is used to characterize the acoustic field near an acoustic source. Higher multipole moments can also be used in the general case. Equation (1) is discretized as

$$\tilde{\mathbf{p}} = \mathbf{G}\mathbf{q} + \mathbf{n}, \quad (4)$$

where $\mathbf{p} = [p(\mathbf{r}_1), p(\mathbf{r}_2), \dots, p(\mathbf{r}_M)]^T \in \mathbb{C}^M$ is the measured pressure at the microphone positions, $\mathbf{q} = [q_1, q_2, \dots, q_N]^T \in \mathbb{C}^N$ is the coefficient vector containing the strength of the sources, which relate to their volume velocity $Q_n = U_n dS$ as $q_i = j\omega\rho Q_i$, and $\mathbf{G} \in \mathbb{C}^{M \times N}$ consists of the free-field Green's function in Eq. (2). The vector \mathbf{n} represents additive Gaussian noise of variance ε . The equivalent sources are generally placed outside the reconstruction domain, retracted away from the surface of the source, to prevent their singularity.

The problem is typically underdetermined ($N \gg M$), as there are more unknown coefficients than measurement points. The classical estimation of \mathbf{q} is typically obtained via the (Tikhonov) regularized Moore-Penrose pseudoinverse

$$\hat{\mathbf{q}} = \mathbf{G}^H(\mathbf{G}\mathbf{G}^H + \lambda\mathbf{I})^{-1}\tilde{\mathbf{p}}, \quad (5)$$

with regularization parameter λ .

From the estimated solution $\hat{\mathbf{q}}$ it is possible to reconstruct the entire sound field elsewhere in the acoustic domain (pressure, velocity, and sound intensity) via the reconstruction matrix \mathbf{G}_s . The matrix consists of the Green's function in Eq. (2) evaluated at the reconstruction points $\mathbf{r} = \mathbf{r}_s$,

$$\mathbf{p}_s = \mathbf{G}_s\hat{\mathbf{q}}, \quad (6)$$

where \mathbf{p}_s is the reconstructed sound pressure. The reconstructed particle velocity (the normal component n) is

$$\mathbf{u}_{n,s} = -\frac{1}{j\omega\rho}[\partial\mathbf{G}_s/\partial n]\hat{\mathbf{q}}, \quad (7)$$

the sound intensity is

$$\mathbf{I}_{n,s} = \frac{1}{2}\text{Re}\{\mathbf{p}_s \otimes \mathbf{u}_{n,s}^*\}, \quad (8)$$

where \otimes is the element-wise (Hadamard) product, and the estimated sound power is $P = \int_S I_n dS$.

B. Least Squares (LS) and Compressive Sensing (CS)

The solution to the system of equations in Eq. (4) can be obtained via an optimization problem, to estimate the unknown equivalent source amplitudes \mathbf{q} . The classical least-squares solution found in Eq. (5) corresponds to the ℓ_2 -norm minimization estimate

$$\hat{\mathbf{q}}_{\text{LS}} = \arg \min_{\mathbf{q}} \|\mathbf{q}\|_2^2 \quad \text{subject to } \|\mathbf{G}\mathbf{q} - \tilde{\mathbf{p}}\|_2 \leq \varepsilon, \quad (9)$$

with the ℓ_p -norm defined as $\|\mathbf{x}\|_p^p = \sum_i |x_i|^p$. Problem (9) seeks the solution $\hat{\mathbf{q}}$ with minimum energy. Alternatively, a sparse solution with CS (Ref. 12) can be obtained by solving the ℓ_1 -norm minimization problem

$$\hat{\mathbf{q}}_{\text{CS}} = \arg \min_{\mathbf{q}} \|\mathbf{q}\|_1 \quad \text{subject to } \|\mathbf{G}\mathbf{q} - \tilde{\mathbf{p}}\|_2 \leq \varepsilon. \quad (10)$$

The CS method makes it possible to represent the observed sound field with a minimal number of waves and it is valid in a wide frequency range.² Nonetheless, the methodology requires that the underlying problem is sparse, which is often problematic in the near-field of a source, or to characterize spatially extended sources.¹²

C. Total Generalized Variation (TGV)

In this study we propose an alternative approach where the problem is formulated in a general sparse analysis framework,²⁰

$$\hat{\mathbf{q}} = \arg \min_{\mathbf{q}} \|\mathbf{D}\mathbf{q}\|_1 \quad \text{subject to } \|\mathbf{G}\mathbf{q} - \tilde{\mathbf{p}}\|_2 \leq \varepsilon, \quad (11)$$

where $\mathbf{D} \in \mathbb{R}^{P \times N}$ constitutes a linear mapping of the coefficients from \mathbb{C}^N to \mathbb{C}^P . The LS and CS solutions in Eqs. (9) and (10) (so-called synthesis priors) impose a direct structure on the reconstructed signal. Contrarily, the generalized approach presented in Eq. (11) (analysis prior) does not

impose a structure directly on the signal, but on a transformed domain instead, and is therefore more general.

The best known analysis sparsity prior is the method of Total Variation (TV) where the matrix \mathbf{D} is a finite difference approximation to the gradient,¹⁹ thus promoting solutions with a sparse gradient. The method has recently been examined for localizing spatially extended acoustic sources.²⁷ The TV solution promotes piecewise constant solutions, consisting of flat regions [where $|\nabla\mathbf{q}(\mathbf{r}_0)| = 0$] and a few occurrences where $|\nabla\mathbf{q}(\mathbf{r}_0)| \neq 0$. The approach is known to result in staircasing effects,²⁷ as the solution cannot conform to continuously varying source amplitudes. Therefore, in general acoustic problems it seems suitable to promote sparsity on the higher-order derivatives, thus promoting solutions that conform to higher-order polynomials.

The second order spatial derivative on the solution yields the cost function

$$\int_{\Omega} |\nabla^2 q(\mathbf{r}_0)| dS, \quad (12)$$

resulting in the second order TGV problem

$$\hat{\mathbf{q}}_{\text{TGV}} = \arg \min_{\mathbf{q}} \|\mathbf{L}\mathbf{q}\|_1 \quad \text{subject to } \|\mathbf{G}\mathbf{q} - \mathbf{p}\|_2 \leq \varepsilon, \quad (13)$$

where \mathbf{L} is an approximation to the two-dimensional Laplacian operator. The approach promotes solutions that are continuous and piecewise linear (with minimal curvature).

D. Fused-TGV

In this study, we propose a fused approach that combines sparsity on the second order derivative and on the solution. The approach promotes block-sparse solutions where the non-zero coefficients are grouped together. The F-TGV solution is

$$\hat{\mathbf{q}}_{\text{FTGV}} = \arg \min_{\mathbf{q}} \left\| \begin{bmatrix} \mathbf{L} \\ \mu\mathbf{I} \end{bmatrix} \mathbf{q} \right\|_1 \quad \text{subject to } \|\mathbf{G}\mathbf{q} - \mathbf{p}\|_2 \leq \varepsilon, \quad (14)$$

where \mathbf{I} is the identity matrix, \mathbf{L} is the second order spatial derivative ∇^2 , and μ is a hyperparameter to balance sparsity on the derivative and on the signal. The matrix \mathbf{D} in Eq. (11) is therefore $\mathbf{D} = [\nabla^2, \mu\mathbf{I}]^T \in \mathbb{C}^{2N \times N}$.

It is possible to motivate the physical significance of this analysis prior. Writing out the objective function of Eq. (14) in continuous form (thus the volume velocity Q is replaced by the normal velocity U_n),

$$\int_{\Omega} |\nabla^2 U_n(\mathbf{r}_0)| + |\mu U_n(\mathbf{r}_0)| dS, \quad (15)$$

which can be seen as promoting block-sparse solutions, and therefore is well-suited to examine both compact as well as extended sources. In fact, for $\mu = k^2$ yields to minimizing a variation integral similar, although not identical, to the Helmholtz equation $(\nabla^2 + k^2)u_n(\mathbf{r})|_{\Omega} = 0$ in the vicinity of the source.

III. NUMERICAL RESULTS

A. Baffled piston

We consider the reconstruction of the sound field radiated by a rigid piston of 20 cm radius with vibrational velocity $1 \mu\text{m/s}$, placed in the xy plane at $z=0$ and centered at $(x, y) = (0, 0)$. A 60-channel microphone array (of 60 cm diameter; microphone positions shown in Fig. 1) is placed at $z_h = 6$ cm. The equivalent sources used to model the source radiation are distributed in a uniform grid of 41×41 over an area of $50 \times 50 \text{ cm}^2$, and retracted 2 cm behind the source ($z_0 = -0.02$ m). Additive noise with 30 dB signal-to-noise (SNR) ratio is added to the simulated pressure measurements. It is assumed that the noise-floor ε can be measured (the influence of a wrong estimation of the noise floor is examined in Sec. III D). Based on the estimated amplitudes of the equivalent sources, the sound field (sound pressure, normal velocity, and sound intensity) is reconstructed at $z_s = 3$ cm. We use the CVX computing package,³¹ based on an interior point method to solve the optimization problem. The F-TGV hyperparameter is set to $\mu = 1$, and for both the TGV and F-TGV methods the second order spatial derivative is estimated from the 5-point finite difference stencil

$$\mathbf{H}_2 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \quad (16)$$

to obtain the \mathbf{L} matrix (second spatial derivative) assuming periodic boundary conditions on the source aperture.

Figure 1 shows the measured sound pressure at 1 kHz and the obtained equivalent source estimates obtained by the five different methods [LS in Eq. (9), CS in Eq. (10), TV as in Eq. (11) with \mathbf{D} as the gradient, TGV as in Eq. (13), and F-TGV as in Eq. (14)]. It can be seen in Fig. 1 that the least-squares estimate recovers the spatial extent of the piston, although the solution is contaminated by side-lobes. The CS solution recovers a sparse estimate of the source that is physically incorrect, as it does not capture the spatial extent of the source. Instead, the piston's vibrational velocity is estimated as the superposition of a few sparse point sources. The TV, TGV, and F-TGV methods recover the spatial extent of the source. TV yields an estimate of constant amplitude (zero gradient) that conforms well to the shape of the piston, with some staircasing effects around the edges. The F-TGV method exhibits a smoother shape with lower sidelobe contamination than LS or TGV due to the ℓ_1 -minimization on the solution, as well as some artifacts at the boundaries, presumably due to the boundary conditions assumed on the solution.

Figure 2 shows the radiated pressure, velocity, and sound intensity fields that are to be reconstructed based on the different estimates of the methods (shown in Fig. 1). For brevity the reconstructed fields are not shown; instead the spatially-averaged relative error is shown in Fig. 3, to illustrate the accuracy of the reconstructed sound field $(\mathbf{p}_s, \mathbf{u}_{n,s}, \mathbf{I}_{n,s})$ based on the different methods. The spatially-averaged error is calculated as

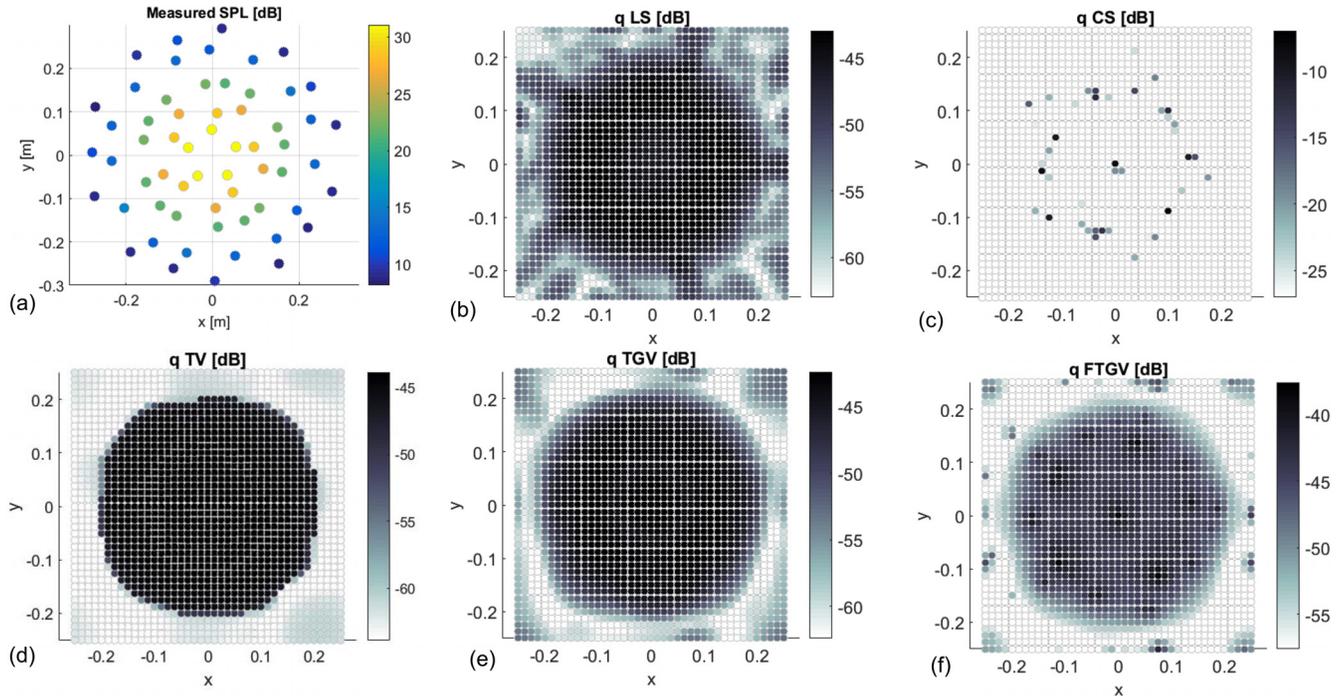


FIG. 1. (Color online) Vibrating piston source. (a) Measured sound pressure at the 60 microphone positions. (b) Estimated coefficients based on LS, (c) CS, (d) TV, (e) TGV, and (f) F-TGV. $f = 1000$ Hz, $z_h = 6$ cm, $z_s = 3$ cm, $z_0 = -2$ cm. All sub-figures have a 20 dB dynamic range.

$$\epsilon = \frac{\|\mathbf{p}_s - \mathbf{p}_{\text{true}}\|}{\|\mathbf{p}_{\text{true}}\|} \times 100, \quad (17)$$

for a frequency range between 100 Hz and 4 kHz (the reconstruction plane is at $z_s = 3$ cm). The figure shows the mean of the error for five occurrences at each frequency.

It is observed in Fig. 3 that the error of the TGV-based methods (TV, TGV, and F-TGV) is lower over a fairly wide frequency band, as they can conform to the extended spatial nature of the source. The CS solution provides a poor estimate of the source radiation, because the source is not spatially sparse, and the sound field reconstruction error is high, particularly at high frequencies. The LS solution captures the spatial extent of the source and provides a good reconstruction at low frequencies, although at higher frequencies the appearance of side-lobes deteriorates the reconstruction, yielding a higher error. The TV method yields an accurate solution at high frequencies, although large errors at low frequencies due to the finite difference operation on the source coefficients, to fit almost constant data. Overall, the TGV and F-TGV methods yield the most accurate results, as the non-zero equivalent sources are grouped on the area of the piston, providing a satisfactory representation of the source over the entire frequency range.

B. Quadrupole

The same study as in Sec. III A is carried out for a longitudinal quadrupole source, with an identical setup as in Sec. III A. The longitudinal quadrupole has a length of $2h = 14$ cm, and it consists of three point sources on a line, with the center source radiating in antiphase with twice the amplitude of the outer two. The equivalent sources are

placed at $z_0 = -5$ cm. Figure 4 shows the error as a function of frequency. It is observable that in this case, the most accurate reconstructions are based on the CS and the F-TGV method. The CS solution favors a sparse representation of the source, which is well-suited to the present example. It is noteworthy that the TV method provides an accurate reconstruction, in particular at high frequencies, where the point sources are well represented by spatially confined areas of constant amplitude. The solution of the F-TGV method is slightly more accurate at low frequencies, as the block-sparse solution is more robust to perturbations and conforms well to the shape of the source in this frequency range. Contrarily, the CS method, due to the high coherence of the sensing matrix at low frequencies,^{6,12} leads to spurious components that do not represent the physical shape of the quadrupole.

Figure 5 shows the estimated coefficients of the quadrupole source by the four methods at 3000 Hz indicating the true point source positions by red crosses. As expected, the results indicate that the LS and TGV methods do not provide a sparse representation of the source, explaining the higher error in the reconstruction. Contrarily, the CS and the F-TGV method yield a sparse solution that coincides with the actual physical source. The TV method provides a solution that exhibits regions of constant amplitude, and which leads to a non-physical representation of the source, although providing an accurate reconstruction of the sound field. It is interesting to note that the block-sparse nature of the F-TGV leads to identifying correctly the positions of the acoustic sources. Both CS and F-TGV promote solutions that are exempt of strong artifacts, respectively, because of promoting direct sparsity in one case, and block-sparsity in the other.

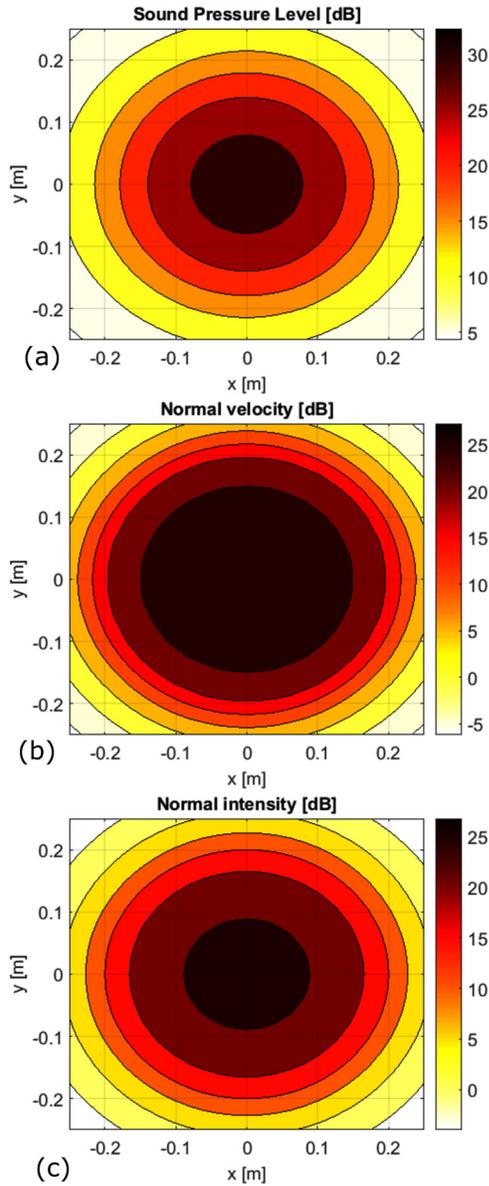


FIG. 2. (Color online) True sound field radiated by the vibrating piston 3 cm in front of it. (a) SPL [dB re 20×10^{-6}], (b) normal component of the particle velocity [dB re 50×10^{-9}], (c) and normal component of the active intensity [dB re 10^{-12}].

C. Spatial response to a point source

To examine the properties of the different solution methods, we examine the spatial impulse response of the methods, i.e., their response to a monopole source (reminiscent of a point-spread function). This analysis provides valuable insight into the solution across frequency, and the capability to obtain a wideband reconstruction. The monopole is placed at $z=0$, and the array at $z_h=8$ cm, and $\text{SNR}=30$ dB. Figure 6 shows the response in the horizontal direction from $x=-0.25$ m to $x=0.25$ m, as a function of frequency (all plots with 25 dB dynamic range for ease of comparison). The LS method shows its characteristic response with a wide main-lobe at low frequencies, which become narrower at high frequencies, where high side-lobes also appear. The CS method recovers nearly an ideal delta function, because it promotes a sparse solution

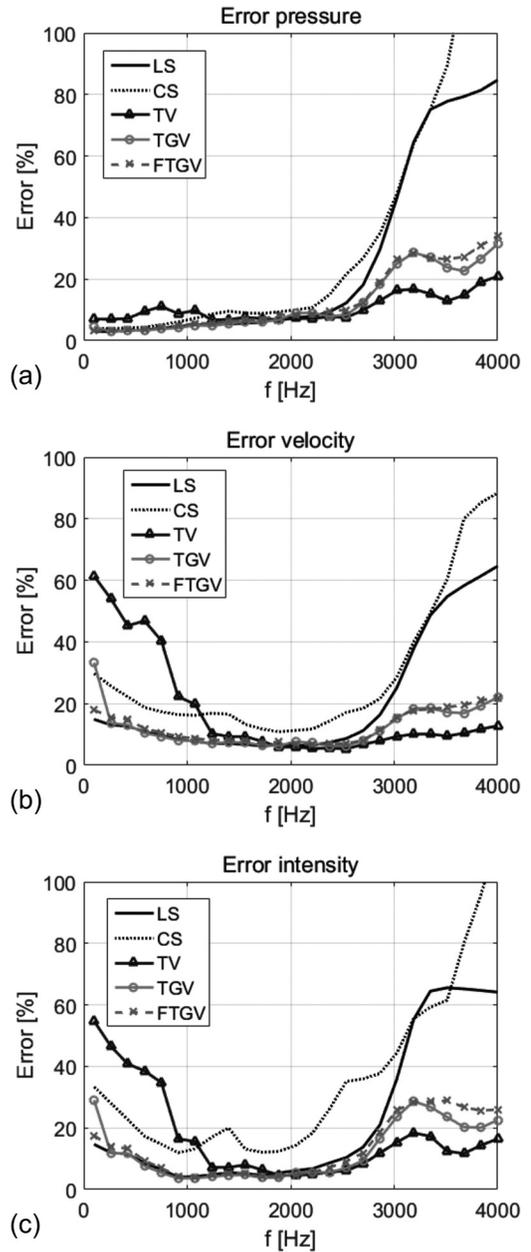


FIG. 3. Reconstruction error, as in Eq. (17), for the reconstructed sound field of a vibrating piston.

directly. The TGV method yields a response which is in fact similar to the LS one, showing a high side-lobe contamination, because the method does not promote spatial sparsity directly. The spatial impulse response of the TV method shows the sensitivity to noise, particularly in the low frequency response, as well as occasional artifacts. In agreement with the previous results, the F-TGV method yields a precise identification of the point source, due to promoting a block-sparse solutions.

D. Noise and noise-floor estimation

Finally we examine the robustness of the methods to perturbations in the measured data. We examine the error due to increasing additive noise, as well as to a wrong estimate of the noise floor. Figure 7 shows the spatially

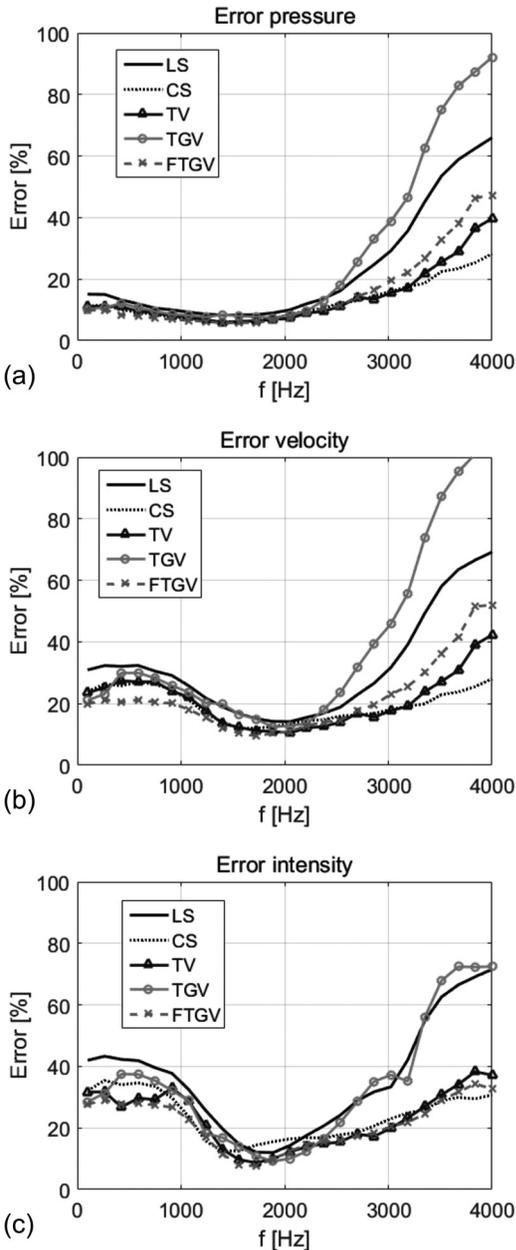


FIG. 4. Reconstruction error, as in Eq. (17), for the reconstructed sound field of a longitudinal quadrupole.

averaged error of the reconstructed sound field (pressure, normal velocity, and intensity), estimated via Eq. (17), for increasing levels of additive white noise. It is still assumed that the noise-floor ε is correctly estimated from the measurement data. The results show the mean error for ten repetitions. The source is a monopole placed at $z = -5$ cm, radiating at 500 Hz, and the array properties and reconstruction settings are as in Sec. III A. The errors show that the TGV method is very sensitive to noise, due to appearance of spurious artifacts. The LS method is generally robust to noise, although in this example the error is not low, because the source is a monopole. CS is fairly robust, as promoting spatial sparsity is effective in suppressing noisy components from the solution (given a correct estimate of the noise-floor); so is the TV regularization, which

is a well-known noise suppressing prior. F-TGV is remarkably robust, especially when estimating particle velocity and sound intensity. Grouping the non-zero elements makes it robust to perturbations, and the sparsity prior tends to suppress spurious components. Similar results are found for other sources (quadrupole, piston), although they are not shown here for brevity.

It is of importance to examine the sensitivity of the methods to a wrong estimation of noise floor. One of the drawbacks of CS is the lack of automatic parameter-choice methods to determine the regularization parameter (unlike regularized least-square inversions). CS relies on either having prior knowledge of the noise floor [where Eq. (10) is suitable], or solving the problem with an *ad hoc* choice of regularization parameter (as with the LASSO, i.e., the unconstrained form of the CS problem).

Figure 8 shows the sensitivity of the methods to a wrong estimation of the relative noise floor (ε). This serves to study the accuracy of the different methods, depending on the prior knowledge on the noise. The source used is the same monopole source as in Fig. 7, with SNR = 30 dB and $f = 500$ Hz. All methods provide a minimum error around $\varepsilon/\varepsilon_{\text{true}} \approx 1$, when $\varepsilon \rightarrow \varepsilon_{\text{true}}$, and show a similar error trend when the solution is over-regularized due to a too high noise estimate. For an underestimation of ε , in agreement with Fig. 7, TGV is the most sensitive method to noise and shows consistently the greatest error. CS and TV are sensitive to the choice of regularization as a slight under-regularization results in spurious sources that affect the reconstruction. LS and F-TGV are the most robust methods, as the under-regularization does not result in spurious components, but rather on a slight mis-estimation of the source amplitude, where the localization of the sound source is not affected.

IV. EXPERIMENTAL RESULTS

An experimental study is conducted to examine the reconstruction by the methods examined in this study. The source is a dipole-like source consisting of two loudspeaker drivers closely mounted face to face, and driven in antiphase. Figure 9 shows the source and the microphone array. The measurements were conducted in the large anechoic chamber (1000 m³) at the Technical University of Denmark, DTU, with a 60-channel Bruel & Kjaer (Naerum, Denmark) *combo microphone array* of 1 m diameter, also shown in Fig. 9. The loudspeakers were driven with white noise, and oriented so that the driver's plane is normal to the array plane. The edge of the source was located at $z = 0$ cm, and the microphone array at $z_h = 16$ cm. A regular grid of 35×35 equivalent sources located in the plane $z_0 = -5$ cm is used to model the source. After estimating the amplitude of the equivalent sources, the sound field is reconstructed on the xy plane at $z_s = 5$ cm.

Figure 10 shows the estimation of the equivalent source amplitudes at 1500 Hz from the LS, CS, TV, and F-TGV methods. The results are consistent with the numerical results. LS features a high level of side-lobes, whereas CS

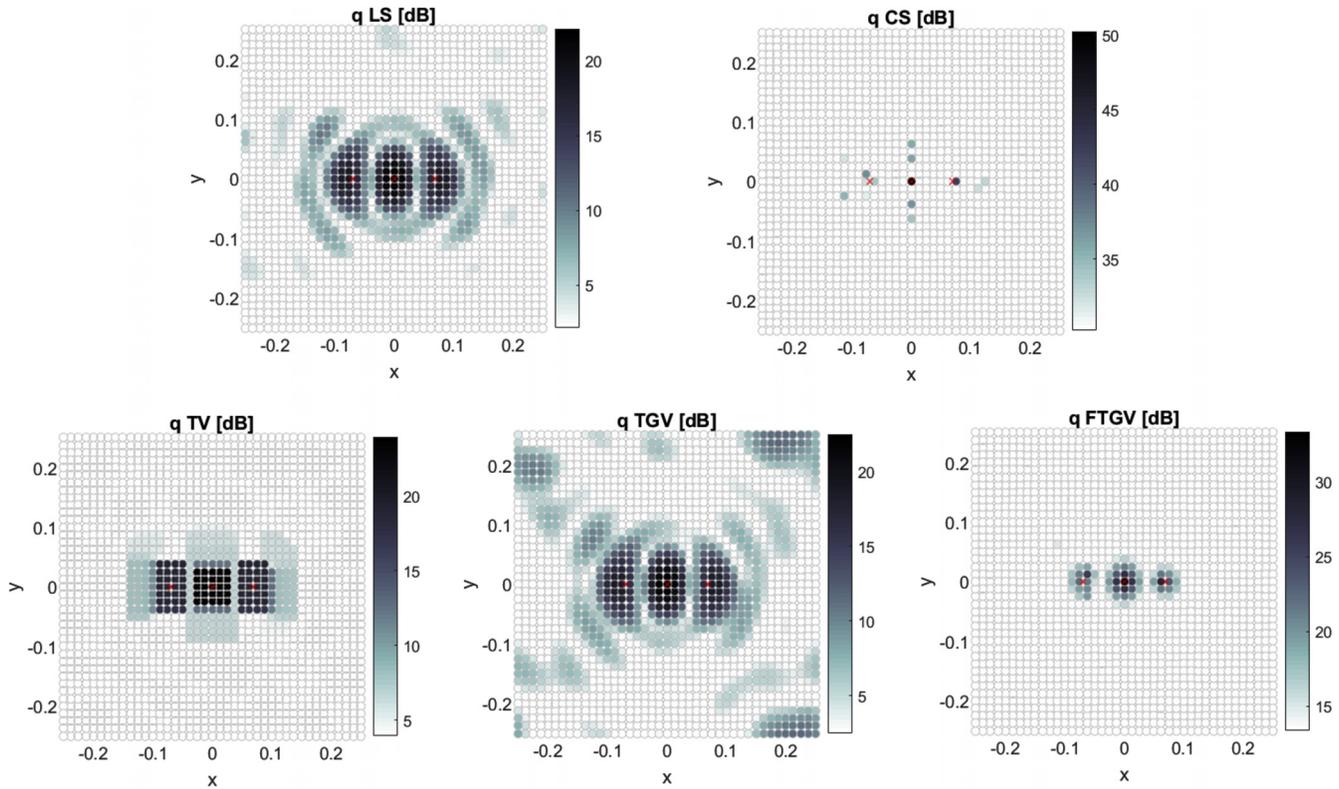


FIG. 5. (Color online) Quadrupole source. Estimated coefficients based on CS, LS, TV, TGV, and F-TGV. $f = 3000$ Hz, $z_h = 6$ cm, $z_0 = -5$ cm. All sub-figures have a 20 dB dynamic range.

yields a sparse representation of the source, with only a few non-zero terms. The TV method exhibits a similar result as in the numerical results, with extended spatial regions of constant amplitude that do not conform to the physical

nature of the source. The proposed F-TGV method yields a block-sparse solution, where the non-zero coefficients are spatially grouped together, resulting in a fairly compact representation of the acoustic source.

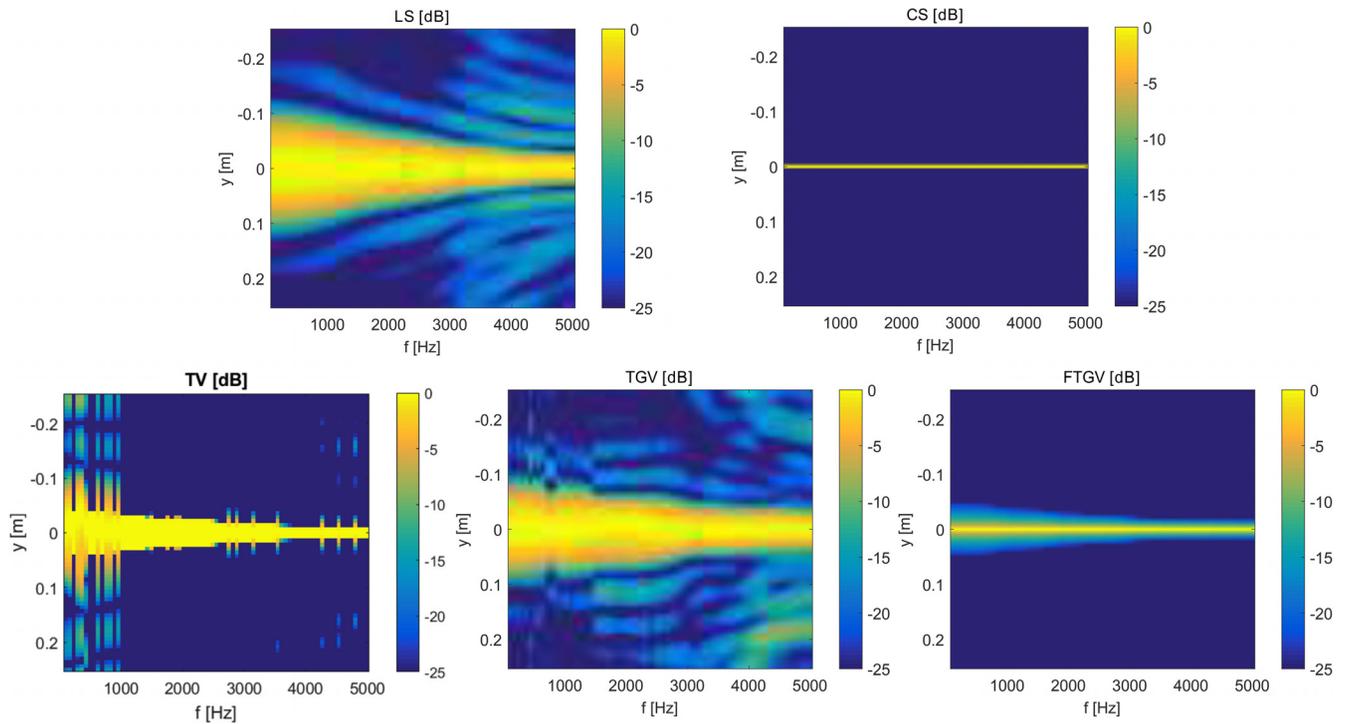


FIG. 6. (Color online) Spatial response of the methods to a monopole located at $(x, y) = (0, 0)$ cm; $z_h = 8$ cm, $z_0 = 0$. Estimated coefficients based on CS, LS, TV, TGV, and F-TGV, SNR = 30 dB.

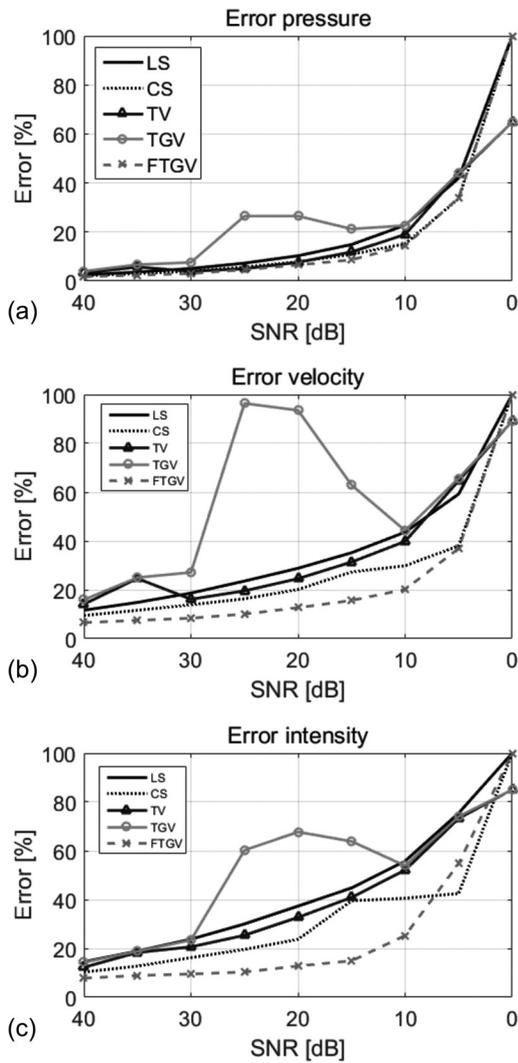


FIG. 7. Error due to increasing additive background noise in reconstructing the sound field radiated by a monopole at 500 Hz. The results display the average error over ten realizations of normally distributed random noise for the methods examined.

Figure 11 shows the reconstructed sound pressure field on the plane $z_s = 5$ cm, based on the LS, CS, and F-TGV methods, at 1500 Hz. Both the CS and F-TGV reconstructions yield a fairly smooth dipolar sound field, with the expected directivity from the source. In the reconstruction with the LS method, there appear spatial artifacts, presumably due to the influence of the side-lobe levels observed in Fig. 10. It is difficult to assess quantitatively the error point by point, mostly due to positioning uncertainty. Therefore we present the averaged sound pressure level (SPL) estimated with the methods on the reconstruction plane, compared to the actual true one measured *in situ*, shown in third-octave bands, at the bottom of Fig. 11. As expected, and in agreement with the numerical results, the LS estimation is somewhat inaccurate at high frequencies, whereas F-TGV, TV, and CS provide a better estimate.

Finally, Fig. 12 shows the estimation of the sound power radiated by the source. The sound power is estimated via reconstruction of the normal intensity at $z_s = 5$ cm, compared to a reference (“true”) sound power. The reference sound

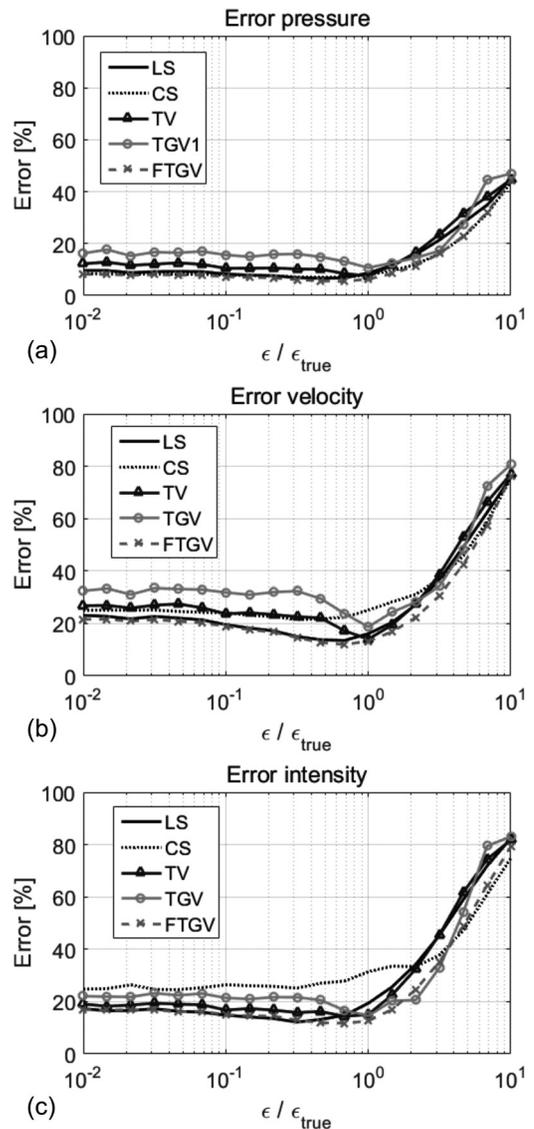


FIG. 8. Error of the methods due to a wrong estimation of the noise floor ϵ .

power is estimated via measurement of the pressure in the reconstruction plane (z_s), to directly reconstruct the sound power in this plane using the ESM (which is subject to error at very high frequencies). The results indicate that the CS, TV, and F-TGV recover the sound power correctly, and

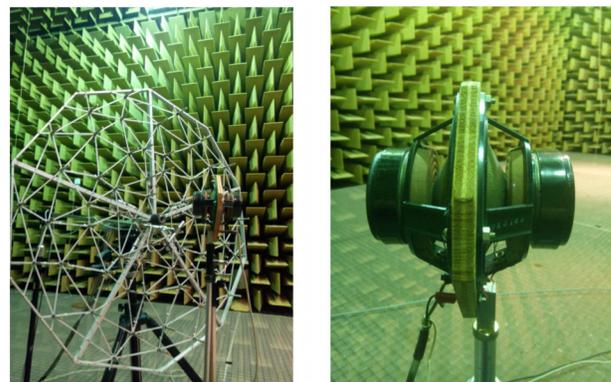


FIG. 9. (Color online) Experimental setup. Microphone array (60-channel) (left). Experimental dipole source (right).

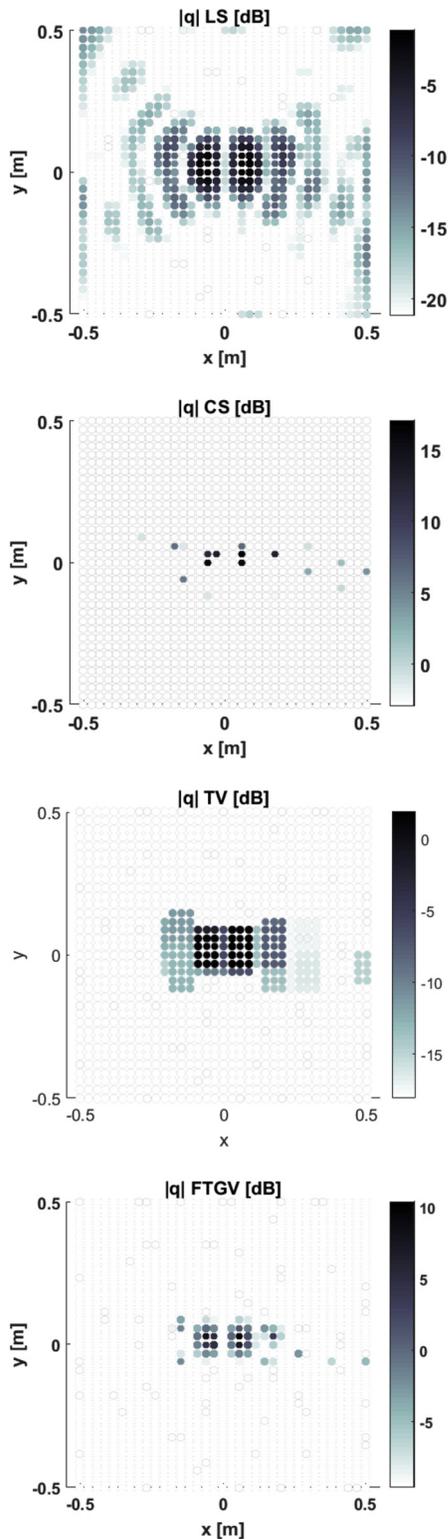


FIG. 10. (Color online) Estimated equivalent source strengths at 1500 Hz.

make it possible to obtain a wideband reconstruction of the radiated sound field, whereas the LS reconstruction tends to underestimate the sound power at high frequencies, as has been identified in previous studies.¹¹ For this source, it seems that CS and F-TGV provide somewhat more accurate quantitative results than the other methods, in agreement with the numerical simulations.

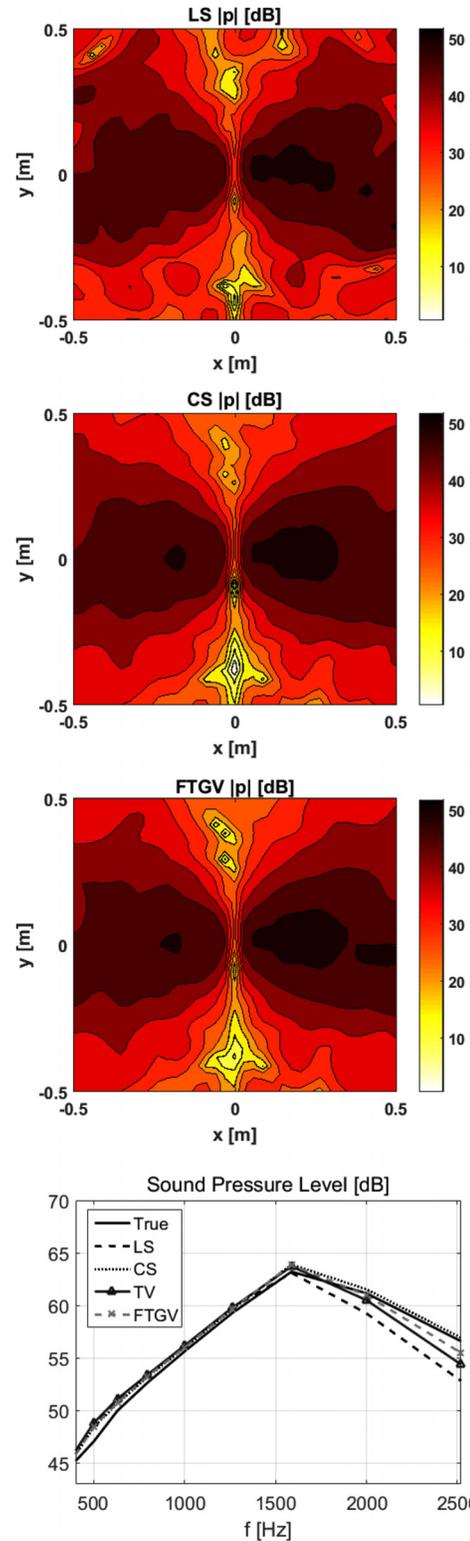


FIG. 11. (Color online) Estimated sound pressures at the reconstruction plane $z_s = 5$ cm, and averaged SPLs compared to the true SPL measured with the microphone array.

V. CONCLUSIONS

This study examines block-sparse methods based on TV approaches for reconstructing acoustic sound fields in the near-field of a source. A F-TGV method has been developed and examined, which promotes block-sparsity by grouping

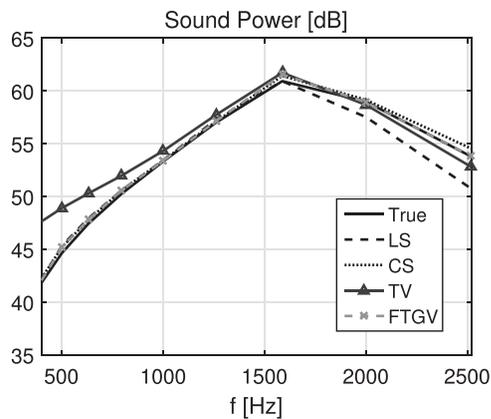


FIG. 12. Estimated sound power levels at the reconstruction plane $z_s = 5$ cm, compared to the true sound power level emitted by the source (estimated via measuring the sound intensity with the microphone array on the reconstruction plane).

the non-zero coefficients. The study shows that both TV and F-TGV can successfully reconstruct spatially extended sources. The numerical and experimental results indicate that the F-TGV method is suitable to examine both compact and extended sources, and expands on the applicability of CS for wideband reconstruction. As a result of promoting block-sparsity, the F-TGV approach is robust to perturbations and to poor prior knowledge of the noise-floor in the measured data. All in all, the approach seems promising for its generality, robustness, and the capability to provide a wideband reconstruction of sound fields that are not necessarily spatially sparse.

On a broader perspective, the results are an encouraging step toward the use of explicit analysis modeling priors, which can lead to general methods that do not necessarily impose a specific structure on the reconstructed sound field, but rather on its underlying physical structure (e.g., block-sparsity in the present study), revealing an interesting potential for the capture and analysis of complex sound fields.

ACKNOWLEDGMENTS

This work was supported by the Danish Council for Independent Research (DFF) under Grant No. 0602-02340B DFF-FTP, and by LABEX WIFI (Laboratory of Excellence within the French Program “Investments for the Future”) under reference ANR-10-IDEX-0001-02 PSL*. We would like to thank Peter Gerstoft and Earl Williams for discussions on the F-TGV method, as well as Niccolò Antonello for an early discussion on TV methods.

- ¹R. G. Baraniuk, “Compressive sensing [lecture notes],” *IEEE Signal Proc. Mag.* **24**(4), 118–121 (2007).
- ²E. J. Candes and M. B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Proc. Mag.* **25**, 21–30 (2008).
- ³M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing* (Springer, New York, 2010), Chap. 1, 376 pp.
- ⁴S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing* (Springer, New York, 2013).

- ⁵G. F. Edelmann and C. F. Gaumont, “Beamforming using compressive sensing,” *J. Acoust. Soc. Am.* **130**(4), EL232–EL237 (2011).
- ⁶A. Xenaki, P. Gerstoft, and K. Mosegaard, “Compressive beamforming,” *J. Acoust. Soc. Am.* **136**(1), 260–271 (2014).
- ⁷P. Gerstoft, A. Xenaki, and C. F. Mecklenbräuker, “Multiple and single snapshot compressive beamforming,” *J. Acoust. Soc. Am.* **138**(4), 2003–2014 (2015).
- ⁸G. N. Lilis, D. Angelosante, and G. B. Giannakis, “Sound field reproduction using the lasso,” *IEEE Trans. Audio, Speech, Lang. Process.* **18**(8), 1902–1912 (2010).
- ⁹W. Jin and W. B. Kleijn, “Theory and design of multizone soundfield reproduction using sparse methods,” *IEEE/ACM Trans. Audio, Speech Lang. Process.* **23**(12), 2343–2355 (2015).
- ¹⁰G. Chardon, L. Daudet, A. Peillot, F. Ollivier, N. Bertin, and R. Gribonval, “Near-field acoustic holography using sparse regularization and compressive sampling principles,” *J. Acoust. Soc. Am.* **132**(3), 1521–1534 (2012).
- ¹¹J. Hald, “Fast wideband acoustical holography,” *J. Acoust. Soc. Am.* **139**(4), 1508–1517 (2016).
- ¹²E. Fernandez-Grande, A. Xenaki, and P. Gerstoft, “A sparse equivalent source method for near-field acoustic holography,” *J. Acoust. Soc. Am.* **141**(1), 532–542 (2017).
- ¹³A. Richard, E. Fernandez-Grande, J. Brunskog, and C.-H. Jeong, “Estimation of surface impedance at oblique incidence based on sparse array processing,” *J. Acoust. Soc. Am.* **141**(6), 4115–4125 (2017).
- ¹⁴E. G. Williams and J. D. Maynard, “Holographic imaging without the wavelength resolution limit,” *Phys. Rev. Lett.* **45**(7), 554–557 (1980).
- ¹⁵J. D. Maynard, E. G. Williams, and Y. Lee, “Nearfield acoustic holography I: Theory of generalized holography and the development of NAH,” *J. Acoust. Soc. Am.* **78**(4), 1395–1413 (1985).
- ¹⁶J. Hald, “Basic theory and properties of statistically optimized near-field acoustical holography,” *J. Acoust. Soc. Am.* **125**(4), 2105–2120 (2009).
- ¹⁷A. Sarkissian, “Method of superposition applied to patch near-field acoustic holography,” *J. Acoust. Soc. Am.* **118**(2), 671–678 (2005).
- ¹⁸C.-X. Bi, Y. Liu, L. Xu, and Y.-B. Zhang, “Sound field reconstruction using compressed modal equivalent point source method,” *J. Acoust. Soc. Am.* **141**(1), 73–79 (2017).
- ¹⁹L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Phys. D* **60**, 259–268 (1992).
- ²⁰S. Vaiter, G. Peyre, C. Dossal, and J. Fadili, “Robust sparse analysis regularization,” *IEEE Trans. Inf. Theory* **59**(4), 2001–2016 (2013).
- ²¹K. Bredies, K. Kunisch, and T. Pock, “Total generalized variation,” *SIAM J. Imaging Sci.* **3**(3), 492–526 (2010).
- ²²R. Tibshirani and J. Taylor, “The solution path of the generalized lasso,” *Annals Stat.* **39**(3), 1335–1371 (2011).
- ²³E. Fernandez-Grande and A. Xenaki, “Compressive sensing with a spherical microphone array,” *J. Acoust. Soc. Am.* **139**(2), EL45–EL49 (2016).
- ²⁴C. R. Vogel and M. E. Oman, “Iterative methods for total variation denoising,” *SIAM J. Sci. Comput.* **17**(1), 227–238 (1996).
- ²⁵A. Chambolle and P.-L. Lions, “Image recovery via total variation minimization and related problems,” *Numerische Mathematik* **76**(2), 167–188 (1997).
- ²⁶A. Beck and M. Teboulle, “Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems,” *IEEE Trans. Image Process.* **18**(11), 2419–2434 (2009).
- ²⁷A. Xenaki, E. Fernandez-Grande, and P. Gerstoft, “Block-sparse beamforming for spatially extended sources in a Bayesian formulation,” *J. Acoust. Soc. Am.* **140**(3), 1828–1838 (2016).
- ²⁸N. P. Vekua, “O Polnote Sistemy Metaharmoniceskikh Funktsii” (“On the completeness of the system of metaharmonic functions”), *Dokl. Akad. Nauk SSSR* **90**, 715–718 (1953) (in Russian).
- ²⁹G. H. Koopmann, L. Song, and J. B. Fahline, “A method for computing acoustic fields based on the principle of wave superposition,” *J. Acoust. Soc. Am.* **86**(6), 2433–2438 (1989).
- ³⁰N. P. Valdivia and E. G. Williams, “Study of the comparison of the methods of equivalent sources and boundary element methods for near-field acoustic holography,” *J. Acoust. Soc. Am.* **120**(6), 3694–3705 (2006).
- ³¹M. Grant and S. Boyd, “Cvx: Matlab software for disciplined convex programming, version 2.0 beta,” Online (2013), <http://cvx.com/cvx> (Last viewed December 2017).