A sparse equivalent source method for near-field acoustic holography

Efren Fernandez-Grande and Angeliki XenakiPeter Gerstoft

Citation: J. Acoust. Soc. Am. **141**, 532 (2017); doi: 10.1121/1.4974047
View online: http://dx.doi.org/10.1121/1.4974047
View Table of Contents: http://asa.scitation.org/toc/jas/141/1
Published by the Acoustical Society of America

A sparse equivalent source method for near-field acoustic holography

Efren Fernandez-Grande[a] and Angeliki Xenaki
Acoustic Technology, Department of Electrical Engineering, Technical University of Denmark, Building 352, Ørsteds Plads, DK-2800 Kongens Lyngby, Denmark

Peter Gerstoft
Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, USA

(Received 16 July 2016; revised 6 December 2016; accepted 16 December 2016; published online 25 January 2017)

This study examines a near-field acoustic holography method consisting of a sparse formulation of the equivalent source method, based on the compressive sensing (CS) framework. The method, denoted Compressive–Equivalent Source Method (C-ESM), encourages spatially sparse solutions (based on the superposition of few waves) that are accurate when the acoustic sources are spatially localized. The importance of obtaining a non-redundant representation, i.e., a sensing matrix with low column coherence, and the inherent ill-conditioning of near-field reconstruction problems is addressed. Numerical and experimental results on a classical guitar and on a highly reactive dipole-like source are presented. C-ESM is valid beyond the conventional sampling limits, making wide-band reconstruction possible. Spatially extended sources can also be addressed with C-ESM, although in this case the obtained solution does not recover the spatial extent of the source.
 [http://dx.doi.org/10.1121/1.4974047]

[EGW] Pages: 532–542

I. INTRODUCTION

Reconstructing and visualizing the sound field near an acoustic source is useful for understanding its sound radiation and identifying the mechanisms that give rise to its acoustic output. Near-field acoustic holography (NAH) is a powerful reconstruction technique, that relies on measuring with an array of microphones in the near-field of the source, in order to reconstruct the entire sound field over a three-dimensional space about the source.[1,2] NAH makes it possible to estimate the sound pressure, particle velocity, and sound intensity vectors in a different position than measured, with enhanced spatial resolution due to capturing the evanescent waves in the near-field of the source. As a consequence, the evanescent waves make the inverse problem severely ill-conditioned. There are numerous NAH methods, which can be based on explicit Fourier transforms,[1–3] plane or spherical wave expansions,[4–6] inverse numerical approaches,[7,8] or other.[9,10] The Equivalent Source Method (ESM), also known as the method of wave superposition, source simulation method, etc.[6,11–13] is a commonly used method in sound radiation and scattering. The method is based on the fundamental idea that an arbitrary wave-field can be expressed as the superposition of waves radiated by a collection of point sources. In the particular case of acoustic holography, the approach is appealing due to its simplicity, computational efficiency, and the potential to reconstruct over non-separable geometries.[6,13]

The ESM, as well as other holography methods, generally give rise to underdetermined problems, since there are often more waves in the model than measurement points. The classical way of solving the problem is in a Least-Squares (LS) sense, by means of a regularized pseudo-inversion. This corresponds to seeking the solution with minimum energy in the solution subspace. However, alternative choices are possible. In acoustics, it is often meaningful to obtain a *sparse* representation of the sound field, i.e., *sparse* in the sense that it consists of few waves or sound sources. Such a solution can be obtained based on the compressive sensing (CS) framework.[14,15] CS is a signal acquisition and processing technique rooted on the idea that data have often a sparse representation in some basis, i.e., they can be described only with few non-zero coefficients.[14,15] This sparse representation can be obtained by means of solving an ℓ_1-norm minimization problem. Sparse approaches have been used in several studies concerned with localizing acoustic sources.[16–23] Recent studies have also examined the use of sparse representations in near-field problems, either by finding a sparse representation in wave number domain using Fourier-based NAH,[24] or direct spatial sparsity.[25–27]

In this paper, we propose a sound field reconstruction method based on the equivalent source method (or method of wave superposition), which makes use of CS for obtaining a sparse representation of the measured wave field in the near-field of an acoustic source. In the following, we refer to the method as Compressive–Equivalent Source Method (C-ESM for brevity). The study addresses the importance of having a non-redundant representation of the observed data, as well as the physical significance of the obtained sparse solutions for the reconstruction of acoustic fields.

[a]Electronic mail: efg@elektro.dtu.dk

II. THE EQUIVALENT SOURCE METHOD

Let us consider an arbitrary sound field in the near-field of an acoustic source, where the sound pressure is sampled in space and time by means of an array of microphones. The sound pressure at a microphone position is expressed as due to a continuum of sources distributed on the surface of the sound source,[11,12,28]

$$p(\mathbf{r}_m,\omega) = \mathrm{j}\omega\rho\int_S q(\mathbf{r}_0)G(\mathbf{r_m},\mathbf{r}_0)dS(\mathbf{r}_0), \tag{1}$$

where $\mathbf{r}_0$ is the position of the equivalent sources, $q(\mathbf{r}_0)$ is their volume velocity, and $\mathbf{r}_m$ the position of a microphone. In practice, the equivalent sources are retracted from the surface of the source (or placed behind it) to prevent the singularity (see Fig. 1). The frequency dependency ω is omitted in the following (the $e^{\mathrm{j}\omega t}$ time convention is adopted).

The function $G(\mathbf{r},\mathbf{r}_0)$ in Eq. (1) corresponds to the Green's function in free-space between an equivalent source at $\mathbf{r}_0$ and a point $\mathbf{r}$ in the sound field reconstruction domain

$$G(\mathbf{r},\mathbf{r}_0) = \frac{e^{-\mathrm{j}k||\mathbf{r}-\mathbf{r}_0||}}{4\pi||\mathbf{r}-\mathbf{r}_0||}, \tag{2}$$

where $||\mathbf{r}-\mathbf{r}_0||$ is the magnitude of the vector difference. It is also possible to express the particle velocity vector $\mathbf{u}(\mathbf{r},\omega)$ from Euler's equation of motion,

$$\mathbf{u}(\mathbf{r},\omega) = -\int_S q(\mathbf{r}_0)\nabla G(\mathbf{r},\mathbf{r}_0)dS(\mathbf{r}_0), \tag{3}$$

where $\nabla G(\mathbf{r},\mathbf{r}_0)$ denotes the gradient of G. It follows from Eqs. (1) and (3) that any arbitrary sound field can be expressed as due to the superposition of the waves radiated by such continuum of point sources (equivalent sources). By inferring the complex amplitudes $q(\mathbf{r}_0)$, it is possible to predict the complete sound field over the source-free domain: pressure, particle velocity, therefore sound intensity and sound power.

In practice, these equations are discretized. Equation (1) becomes

$$p(\mathbf{r}_m) = \sum_{n=1}^{N} q_n G(\mathbf{r}_m,\mathbf{r}_{0,n}). \tag{4}$$

The problem reduces to a simple system of linear equations, in matrix form,

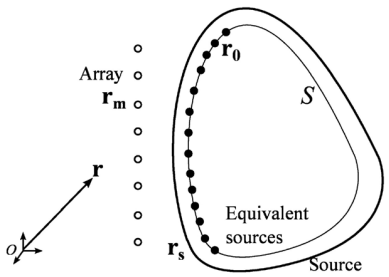


FIG. 1. Equivalent source diagram. Note that the equivalent sources are retracted behind the source (away from the reconstruction area), to avoid their singularity being in the field.

$$\mathbf{p} = \mathbf{G}\mathbf{q}, \tag{5}$$

where $\mathbf{p} = [p(\mathbf{r}_1,\omega),...,p(\mathbf{r}_M,\omega)]^T \in \mathbb{C}^M$ is the measured pressure at M microphones, $\mathbf{q} = [q_1,...,q_N]^T \in \mathbb{C}^N$ is the coefficient vector containing the strength of the N sources, which relate to their volume velocity Q_n as $q_n = \mathrm{j}\omega\rho Q_n$, and $\mathbf{G} \in \mathbb{C}^{M\times N}$ contains the entries from the free-field Green's function describing the propagation between the positions of the equivalent sources $\mathbf{r}_{0,n}$ and microphones $\mathbf{r}_m$.

The solution of Eq. (5) leads to an estimate $\hat{\mathbf{q}}$, which makes it possible to reconstruct the sound field. It is thus possible, as shown in Eqs. (1)–(4) to reconstruct the entire sound field, via a reconstruction matrix $\mathbf{G}_s$, its elements consisting of the Green's function $G(\mathbf{r_s},\mathbf{r}_0)$ from the equivalent sources to the reconstruction points $\mathbf{r}_s$,

$$\mathbf{p_s} = \mathbf{G_s}\hat{\mathbf{q}}, \tag{6}$$

$$\mathbf{u_s} = -\frac{1}{\mathrm{j}\omega\rho}\mathbf{G}_s^n\hat{\mathbf{q}}, \tag{7}$$

where $\mathbf{G}_s^n = \partial\mathbf{G}_s/\partial n$ is the n-component of the gradient of $G(\mathbf{r}_{s,i},\mathbf{r}_{0,j})$. Once the sound pressure and particle velocity are estimated, the sound intensity vector $\mathbf{I} = \mathrm{Re}\{p\,\mathbf{u}^*\}/2$ and sound power $P = \int_S \mathbf{I}\cdot\mathrm{d}\mathbf{S}$ can be obtained.

The estimation of the source strengths $\hat{\mathbf{q}}$ is crucial for the accuracy of the reconstruction, because the problem is severely ill-conditioned: the free-field Green's function models the decay with distance of the spherical waves, and its inversion entails an amplification that should be regularized. Additionally, the problem is typically underdetermined $N > M$, i.e., more coefficients than measurement points, leading to a non-unique solution subspace.

III. METHODOLOGY (C-ESM)

Solving the underdetermined system of equations (5) for the unknown coefficient vector $\mathbf{q}$ requires regularization to constrain the coefficient vector towards desirable (meaningful) solutions. Commonly, from all vectors $\mathbf{q}$ which satisfy Eq. (5), we seek the one with the minimum ℓ_x-norm defined as,

$$||\mathbf{q}||_x = \left(\sum_{i=1}^{N}|q_i|^x\right)^{1/x}. \tag{8}$$

The estimation constitutes an optimization problem which can be formulated as,

$$\hat{\mathbf{q}}(\varepsilon) = \arg\min_{\mathbf{q}} ||\mathbf{q}||_x^x \quad \text{subject to } ||\tilde{\mathbf{p}} - \mathbf{G}\mathbf{q}||_2 \le \varepsilon, \tag{9}$$

where $\varepsilon \ge \|\mathbf{n}\|$ is the estimated noise floor.

An alternative formulation leads to the regularization problem,

$$\hat{\mathbf{q}}(\lambda) = \arg\min_{\mathbf{q}} \|\tilde{\mathbf{p}} - \mathbf{G}\mathbf{q}\|_2^2 + \lambda\|\mathbf{q}\|_x^x, \tag{10}$$

where the regularization parameter $\lambda > 0$ controls the relative importance between the data fitting term and the ℓ_x-norm of the solution vector. For some values of ε and λ, the

constrained formulation [Eq. (9)] and unconstrained formulation [Eq. (10)] are equivalent. The selection of the regularisation parameter λ in (10) requires an iterative solution of the problem along with prior knowledge on the relative noise level.[20] In this study, we make use of Eq. (9), as it is appropriate for applications where the maximum noise level in the data is known or can be estimated.

The "classical" approach uses the ℓ_2-norm to promote smooth, minimum-energy estimates through the convex problem (resulting in the conventional ESM estimation),

$$\hat{\mathbf{q}}_{\ell_2}(\varepsilon) = \arg\min_{\mathbf{q}} ||\mathbf{q}||_2^2 \quad \text{subject to} ||\tilde{\mathbf{p}} - \mathbf{G}\mathbf{q}||_2 \leq \varepsilon. \tag{11}$$

Note that the standard form of the analytic solution to the ℓ_2-norm regularized least-squares problem, $\hat{\mathbf{q}}(\lambda) = \mathbf{G}^H(\mathbf{G}\mathbf{G}^H + \lambda\mathbf{I})^{-1}\mathbf{p}$, is derived from Eq. (10).

Sparse solutions are obtained ideally by solving the ℓ_0-*pseudo norm* problem, where $\|\mathbf{q}\|_0 = (i|q_i \neq 0)$ is a count of the non-zero terms in the vector $\mathbf{q}$. However, this problem constitutes a combinatorial search that is non-convex and easily becomes intractable. On the contrary, CS postulates that the ℓ_0-norm minimization problem can be relaxed to a convex ℓ_1-norm minimization problem when the underlying problem is sparse and the columns of the sensing matrix [$\mathbf{G}$, in the problem of Eq. (5)] are sufficiently incoherent (see Sec. III A). Hence, solving the convex problem,

$$\hat{\mathbf{q}}_{\ell_1}(\varepsilon) = \arg\min_{\mathbf{q}} ||\mathbf{q}||_1 \quad \text{subject to} ||\tilde{\mathbf{p}} - \mathbf{G}\mathbf{q}||_2 \leq \varepsilon, \tag{12}$$

promotes sparse solutions. The C-ESM method makes use of this sparse estimate $\hat{\mathbf{q}}_{\ell_1}(\varepsilon)$. In this study, we use the CVX computing package,[32] based on an interior-point method, which is robust for sparse recovery problems.[33] Although this method is suitable for the size of the problems considered here, when the size of the optimization problem increases significantly (i.e., large and/or dense equivalent source grids), the approach becomes slow. For large problems, it may be of interest to employ alternative algorithms[34] as, e.g., Refs. 35 and 30 (FISTA) and 36 (Sparse Bayesian learning).

A. Sensing matrix coherence in NAH/ESM

A sensing matrix $\mathbf{G}$ with incoherent columns will guarantee a perfect recovery. However, due to the underlying physical structure of the problem, coherence in $\mathbf{G}$ will affect the robustness of the reconstruction, for closely spaced source positions and in the presence of noise.

A general coherence measure for the sensing matrix $\mathbf{G}$ is the restricted isometry property (RIP) condition,[29]

$$(1 - \delta_l)||\mathbf{q}||_2^2 \leq ||\mathbf{G}\mathbf{q}||_2^2 \leq (1 + \delta_l)||\mathbf{q}||_2^2, \tag{13}$$

where $\delta_l \geq 0$ is defined as the lth restricted isometry constant of the matrix $\mathbf{G}$ for an l-sparse vector $\mathbf{q}$. Define $\mathbf{G}_L$ the submatrix composed by any subset L of maximally l normalized columns of $\mathbf{G}$. Then, the condition (13) implies that the Gramian matrix $\mathbf{\Gamma}_L = \mathbf{G}_L^H\mathbf{G}_L$ has its eigenvalues in the interval $[1 - \delta_l, 1 + \delta_l]$. Thus, for $\delta_l < 1$ the Gramian of $\mathbf{G}_L$ has full rank and the matrix $\mathbf{G}$ satisfies the RIP of order l. It follows that the sequence of restricted isometry constants is non-decreasing, $\delta_l \leq \delta_{l+1}$.[29]

The simplest coherence measure is the mutual coherence, which describes the maximum linear dependency between any two different column-vectors of $\mathbf{G} = [\mathbf{g}_1, ..., \mathbf{g}_N]$,

$$\mu(\mathbf{G}) = \max_{1\leq i\neq j\leq N} \Gamma_{ij} = \max_{1\leq i\neq j\leq N} \frac{||\mathbf{g}_i^H\mathbf{g}_j||}{||\mathbf{g}_i|| \cdot ||\mathbf{g}_j||}, \tag{14}$$

where the matrix $\mathbf{\Gamma} = [\Gamma_{ij}]$ is the Gramian of $\mathbf{G}$. It follows from the RIP condition that $\mu = \delta_2 \leq \delta_{l>2}$ thus the mutual coherence is a sufficient indicator of coherence.

An upper bound of the condition number of the sensing matrix $\mathbf{G}$, i.e., the ratio between its largest and smallest singular value, is given by the mutual coherence as,

$$\text{cond}(\mathbf{G}) \leq \frac{\sqrt{1 + \mu(\mathbf{G})}}{\sqrt{1 - \mu(\mathbf{G})}}. \tag{15}$$

This follows from the fact that the singular values of $\mathbf{G}$ are equal to the square root of the eigenvalues of its Gram matrix $\mathbf{\Gamma}$ which are in the interval $[1 - \delta_2, 1 + \delta_2]$. Hence, the closer $\mu(\mathbf{G})$ is to unity, the more ill-conditioned is the matrix $\mathbf{G}$. The condition number is an indication of how well-invertible a matrix is. Although here we do not perform an explicit inversion, the condition number is still representative of the underlying physical problem (i.e., that the waves decay as they propagate away from the source, therefore can pose a challenge in the back-propagation process).

For the matrix $\mathbf{G}$ with elements defined in Eq. (2) the mutual coherence is,

$$\begin{aligned}\mu(\mathbf{G}) &= \max_{1\leq i\neq j\leq N} \frac{\left|\sum_{m=1}^{M} \frac{e^{jk(\|\mathbf{r}_m - \mathbf{r}_{0i}\| - \|\mathbf{r}_m - \mathbf{r}_{0j}\|)}}{\|\mathbf{r}_m - \mathbf{r}_{0i}\|\|\mathbf{r}_m - \mathbf{r}_{0j}\|}\right|}{\sqrt{\sum_{m=1}^{M} \frac{1}{\|\mathbf{r}_m - \mathbf{r}_{0i}\|^2}}\sqrt{\sum_{m=1}^{M} \frac{1}{\|\mathbf{r}_m - \mathbf{r}_{0j}\|^2}}} \\ &= \max_{1\leq i\neq j\leq N} \frac{\left|\sum_{m=1}^{M} \frac{e^{-jk\Delta_j}}{\|\mathbf{r}_m - \mathbf{r}_{0i}\|^2 + \Delta_j\|\mathbf{r}_m - \mathbf{r}_{0i}\|}\right|}{\sqrt{\sum_{m=1}^{M} \frac{1}{\|\mathbf{r}_m - \mathbf{r}_{0i}\|^2}}\sqrt{\sum_{m=1}^{M} \frac{1}{(\|\mathbf{r}_m - \mathbf{r}_{0i}\| + \Delta_j)^2}}},\end{aligned} \tag{16}$$

where $\|\mathbf{r}_m - \mathbf{r}_{0j}\| = \|\mathbf{r}_m - \mathbf{r}_{0i}\| + \Delta_j$. Hence, Eq. (16) shows that as $\Delta_j \to 0$, i.e., $\|\mathbf{r}_m - \mathbf{r}_{0i}\| \approx \|\mathbf{r}_m - \mathbf{r}_{0j}\|$, the mutual coherence approaches unity, $\mu(\mathbf{G}) \to 1$, resulting in a severely ill-conditioned problem. The coherence of $\mathbf{G}$ increases when dense grids are required to achieve fine resolution reconstruction and/or when the distance between the measurement plane and the source plane is increased, as depicted in the schematic of Fig. 2. For large standoff distances ($||\mathbf{r}_{0i} - \mathbf{r}_{0j}||$) or dense equivalent source meshes, the mutual coherence increases as shown in Eq. (16), resulting in estimation error. Contrarily if the equivalent source mesh is less dense, the coherence will decrease, but the equivalent

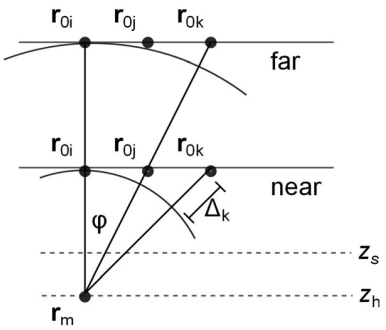
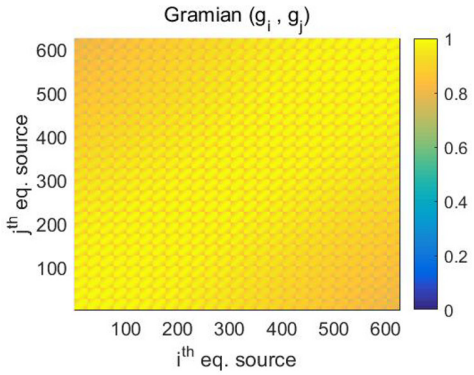


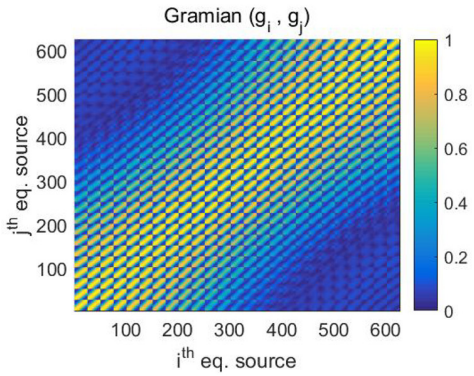
FIG. 2. Schematic showing the relative distance between a measurement point $\mathbf{r}_m$ and a source point $\mathbf{r}_0$ for several source positions i, j, k both in a near and a far plane from the array. The difference in relative distances, $\Delta_k/||\mathbf{r}_m - \mathbf{r}_{0i}||$, decreases for nearby source positions in the same plane or for the same source position at a distant plane.

source model cannot model rapid spatial changes of the sound field.

Figure 3 shows an example of the Gramian matrix [$\mathbf{\Gamma} \in \mathbb{R}^{N\times N}$, with elements Γ_{ij} defined as in Eq. (14)] of the sensing matrix $\mathbf{G}$ in Eq. (5). We consider a random array of 60 microphones and 1 m diameter (see Sec. V). The equivalent source grid consists of 25×25 sources uniformly distributed over an area of 1×1 m^2, and situated at 12 cm distance from the array, in a parallel plane. The apparent block pattern (Fig. 3) results from the vectorization of the equivalent source grid.



(a) 250 Hz



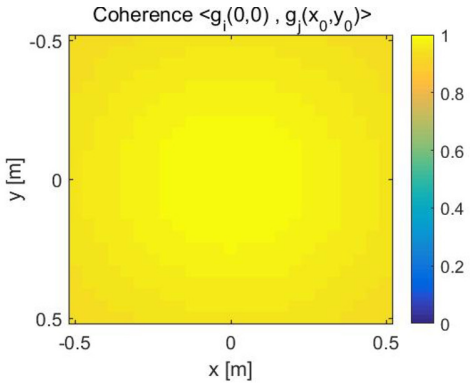
(b) 1200 Hz

FIG. 3. (Color online) Gramian for the sensing matrix $\mathbf{G}$ of size 625×625 constructed from a grid of 25×25 equivalent sources and a 1 m diameter 60 channel microphone array.

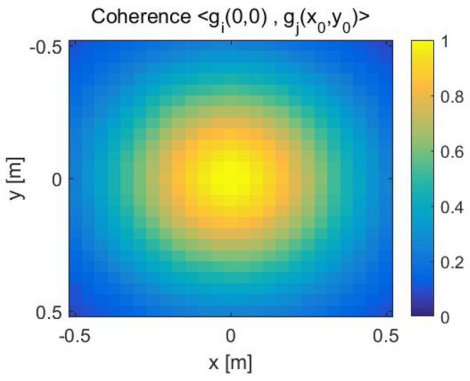
Because of the underlying physical structure of the problem, the column-vectors of $\mathbf{G}$ which correspond to neighboring source locations are highly correlated, as apparent from the Gramian in Fig. 3. At low frequencies, even the column-vectors of $\mathbf{G}$ corresponding to distant equivalent source locations become correlated [Fig. 3(a)], whereas at higher frequencies [Fig. 3(b), 1200 Hz], this is only the case for neighboring equivalent sources. To relate the coherence structure with the equivalent source positions for this measurement, Fig. 4 shows a row of the Gramian matrix with its values rearranged to the source grid. The middle row is chosen (row $i = 313$) to show the correlation of the column-vectors of $\mathbf{G}$ due to the equivalent source at the center of the grid. The higher is the correlation between the column-vectors of $\mathbf{G}$, the more susceptible is the accuracy of source localization to noise. The high coherence lobe indicates the possible extent of erroneous localization.[19]

B. Sparsity in ESM

If the problem under study is inherently sparse, CS guarantees perfect recovery provided that the mutual coherence of the sensing matrix (see Sec. III A) is sufficiently low. In sound source localization problems, the sound field is often due to sources that are confined to a specific location, and most of the solid angle from the array is source-free. In the particular case of near-field problems, the source is sparse when it consists of localized radiation "hot-spots" (or when



(a) 250 Hz



(b) 1200 Hz

FIG. 4. (Color online) Coherence between the center of the equivalent source grid $q_{313}(0,0)$, and the other positions in the grid $q(x, y)$.

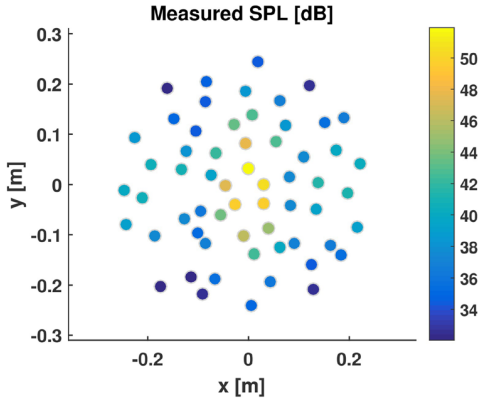
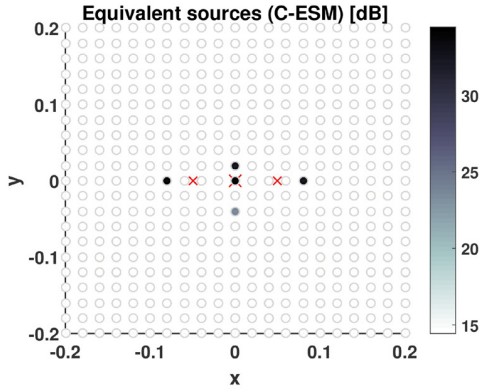
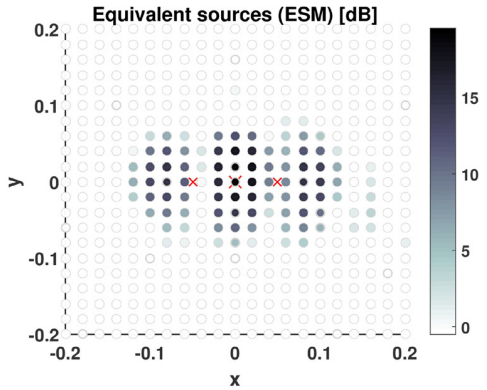


FIG. 5. (Color online) Measured sound pressure level due to a longitudinal quadrupole on the x-y plane, at a distance $z_h = 10\,\text{cm}$ from the array. The sound pressure level (SPL) is plotted at the microphone positions, indicating the array geometry.

the radiated field can be described by a few functions). However, it is also frequent to have spatially extended sources that cover a large solid angle from the array and are not sparse. It is therefore of importance to consider when a sound source consists of localized or extended sources (being inherently sparse or not), as the latter case affects the accuracy of the sound field extrapolation in Eqs. (6) and (7). This is illustrated experimentally in Secs. IV and V.



(a) C-ESM



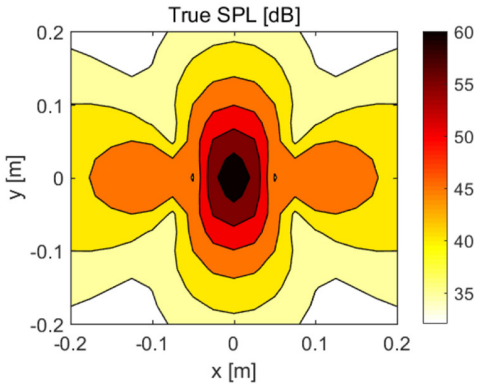
(b) ESM

FIG. 6. (Color online) Equivalent source strengths, i.e., solution to the system of Eq. (5), $z_0 = 0$. (a) Solution based on the proposed C-ESM method—Eq. (12). (b) ESM as in Eq. (11). True point source positions indicated with $\times$.

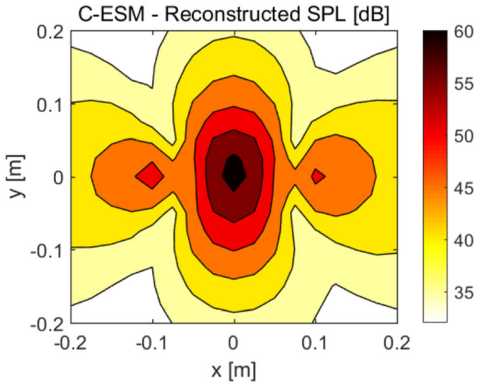
IV. NUMERICAL RESULTS

A simulation is conducted to examine the C-ESM method. The estimated solution vector $\hat{\mathbf{q}}$ and the reconstructed sound field ($\mathbf{p}$ and $\mathbf{u}$) are analyzed, comparing C-ESM [based on Eq. (12)], with the conventional ESM solution [based on (11)].

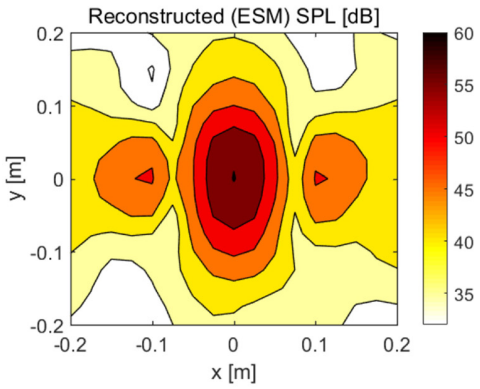
Let us consider a longitudinal quadrupole, of length 10 cm, radiating at $f = 500\,\text{Hz}$. A 50 cm diameter 60 channel microphone array, with pseudo-random spatial sampling is used. The quadrupole is placed 10 cm away from the array ($z_h = 10\,\text{cm}$) and centered on it. The reconstruction takes



(a) True



(b) C-ESM



(c) ESM

FIG. 7. (Color online) Reconstructed sound pressure level of a longitudinal quadrupole. Reconstruction at $z_s = 6\,\text{cm}$. (a) True pressure. (b) C-ESM reconstruction. (c) ESM reconstruction.

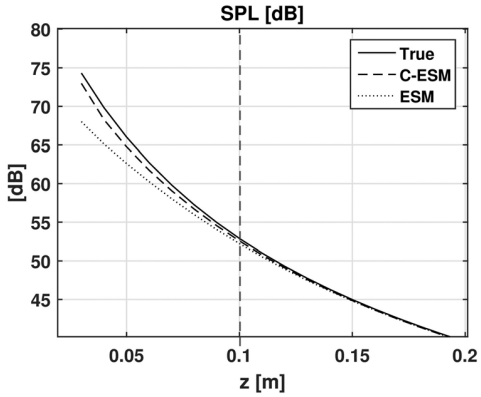


FIG. 8. True and reconstructed sound pressure level as a function of distance for a quadrupole, as in Fig. 5. The vertical dashed line denotes the measurement plane.

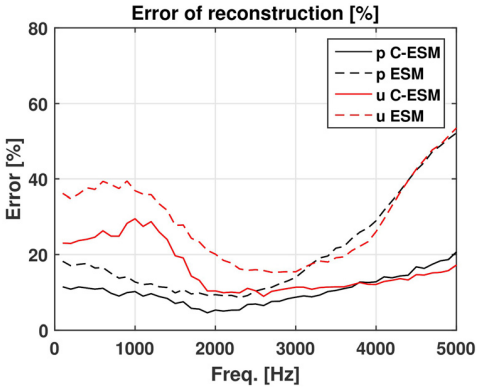


FIG. 10. (Color online) Spatially averaged error, Eq. (18), of the sound pressure reconstruction (black lines) and particle velocity reconstruction (red lines) versus frequency.

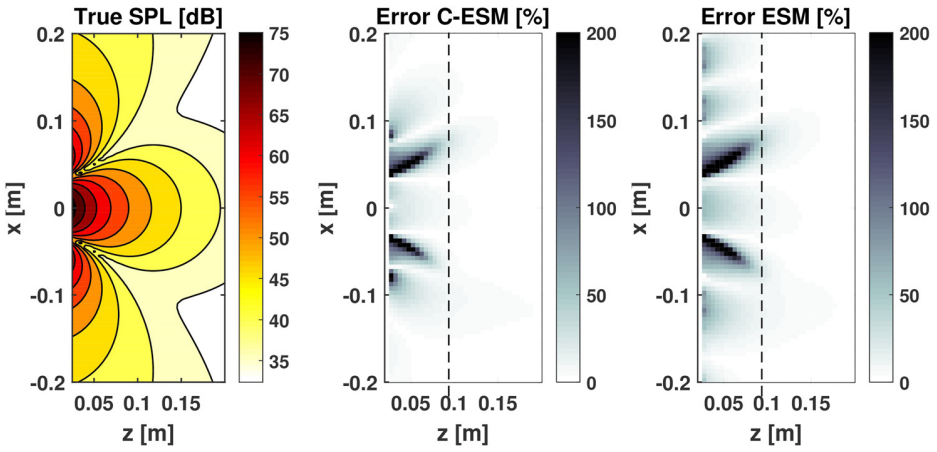
place at 6 cm away from the source ($z_s = 6$ cm), thus entailing a back-propagation of the wave field. The equivalent sources, consisting a grid of 21×21 sources, are placed on the same plane as the quadrupole. Complex Gaussian noise is added to the pressure with a 30 dB spatially averaged signal-to-noise ratio, SNR, defined as $\mathrm{SNR} = 20\log(||\mathbf{p}||/||\mathbf{n}||)$. Figure 5 shows the simulated measured field. The problem is solved as in Eq. (12), estimating a noise floor $\varepsilon = ||\mathbf{p}|| \cdot 10^{-\mathrm{SNR}/20}$, assuming that the SNR is known.

Figure 6 shows the coefficient solution vector, i.e., the recovered equivalent source distribution, for both the proposed C-ESM and the conventional ESM solution (least squares). In the C-ESM case, the solution is sparse, and the location of the quadrupole is visible (the location of the equivalent sources are slightly different than the true ones, presumably due to local high coherence of the columns of $\mathbf{G}$). The ESM solution recovers a source distribution that is far from sparse, and has minimum energy, as required by Eq. (11). It is observable that $||\hat{\mathbf{q}}_{\ell_2}||_2 = 1.9 \times 10^{-6}$, $||\hat{\mathbf{q}}_{\ell_1}||_2 = 4.03 \times 10^{-6}$, and that $||\hat{\mathbf{q}}_{\ell_2}||_1 = 22.1 \times 10^{-6}$ while $||\hat{\mathbf{q}}_{\ell_1}||_1 = 5.9 \times 10^{-6}$. This indicates that the power of the C-ESM solution is greater than ESM, although the sum of all coefficient magnitudes is lower, as expected from the objective functions defined in (11) and (12).

There is an analogy between the recovered solution shown in Fig. 6 and sound source localization methods (where the conventional least squares solution corresponds to a convolution between the ideal source map and the array response, yielding a "blurred" image of the true solution).[30] The apparent source mislocation in Fig. 6(a) can be understood from Sec. III A, considering the high coherence lobe in the source map, and the sensitivity to measurement noise.

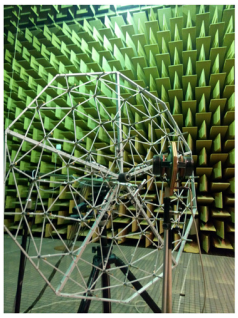
Figure 7 shows the true and reconstructed pressures with C-ESM and the conventional ESM on the reconstruction plane at $z_s = 6$ cm. The spatially averaged error of the reconstruction $\epsilon_p = ||\mathbf{p}_t - \tilde{\mathbf{p}}||/||\mathbf{p}_t||$ is of 10%, indicating a fairly accurate reconstruction. In the case of ESM, ϵ_p is 17%, and deviations near the edges are observable.

Figure 8 shows the sound pressure level along a line on the array axis $(0, 0, \mathbf{z})$, from $z = 3$ cm until $z = 20$ cm at points spaced $\Delta z = 1$ cm, for both the C-ESM method and the ESM. Both methods are accurate near the measurement aperture, where the data are fitted [the $||\tilde{\mathbf{p}} - \mathbf{G}\mathbf{q}||_2 \leq \varepsilon$ constrain in Eq. (9)], as well as in the domain $z_h > z_s$, where the reconstruction constitutes a forward problem. However, closer to the source $(0 < z < z_h)$, where the reconstruction is an inverse problem, we can see differences between the methods. The estimation with the C-ESM method is more accurate, because the obtained coefficients approximate better the acoustic source, as apparent from Fig. 6.

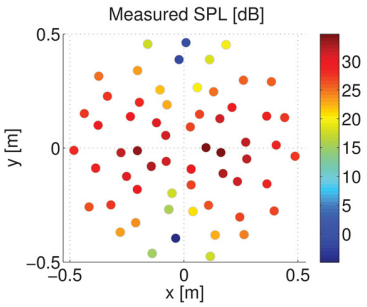


(a) True (b) C-ESM (c) ESM

FIG. 9. (Color online) Relative error [Eq. (17)] of the sound pressure reconstruction near the array (the dashed line indicates the array measurement plane $z_h = 10$ cm). (a) True SPL; (b) Error of C-ESM method; (c) Error of ESM.



(a) Setup



(b) Measured SPL

FIG. 11. (Color online) (a) Measurements in DTU's anechoic chamber, where the dipole source is visible. (b) Measured sound pressure level at 250 Hz.

The reconstruction error is further examined in Figs. 9 and 10. From a true pressure $\mathbf{p}$ and a reconstructed pressure $\tilde{\mathbf{p}}$, the relative error of each element is

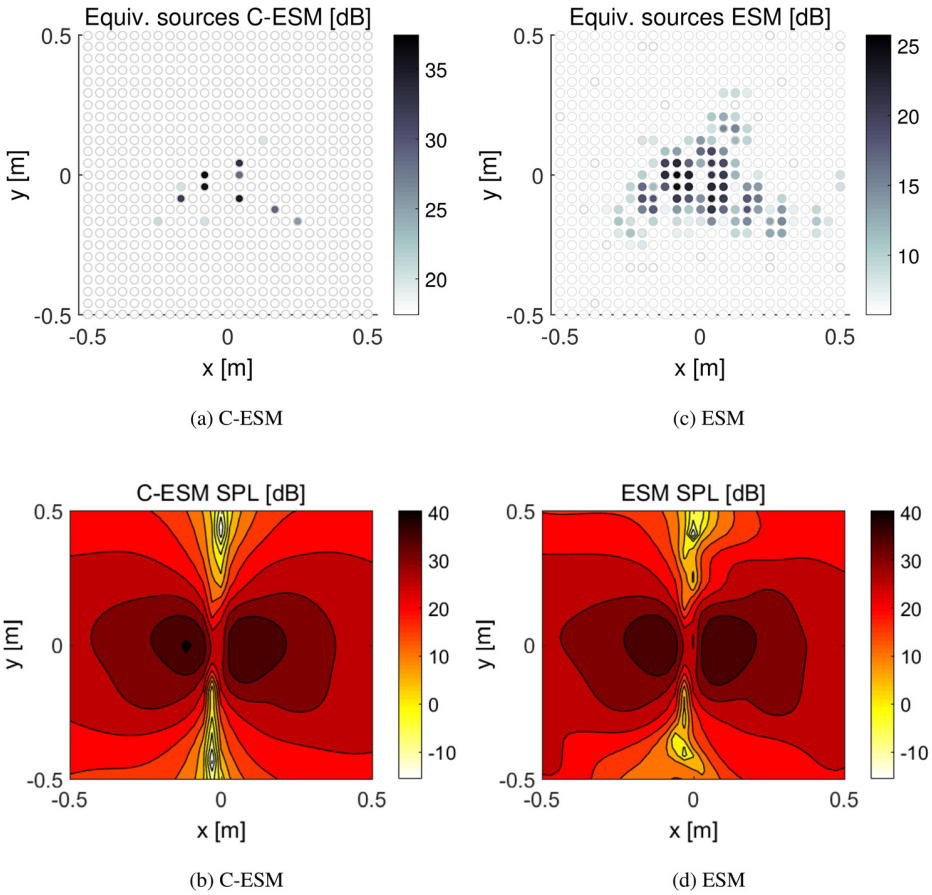
$$e_i = \frac{||p_i - \tilde{p}_i||}{||p_i||}, \tag{17}$$

and the spatially averaged relative error (normalized root-mean-square error) is given as

$$\epsilon = \left(\sqrt{\sum_i (p_i - \tilde{p}_i)^2}\right) \Big/ \left(\sqrt{\sum_i p_i^2}\right). \tag{18}$$

Figure 9 shows the true quadrupole field, and the error of the sound pressure reconstruction for ESM and C-ESM methods. The error is generally higher in the areas where there is a rapid change of the sound field, or values of the sound pressure close to zero. It is remarkable that the C-ESM method is more accurate in the inverse region, i.e., $z_s < z_h$ (z_s is the reconstruction plane and z_h the measurement plane), and in particular, in the area far from the center of the array. This is due to the fact that the CS solution provides a more faithful representation of the actual acoustic source, and the wave extrapolation used for the reconstruction is more accurate than the least square estimate (ESM). These results are consistent with the ones shown in Fig. 8.

The spatially averaged error over the reconstruction area [Eq. (18)] versus frequency is shown in Fig. 10. The simulation setup is the same (except that the equivalent source grid is now 17×17). It is apparent from Fig. 10 that the C-ESM solution is accurate over a broader frequency range, due to the ℓ_1-norm minimization. In the case of conventional ESM, the random microphone array prevents a sharp aliasing limit (above which the reconstruction would be meaningless) at the cost of having greater sidelobe levels that result in higher error, as observed in Fig. 10. When solving the problem in a least-squares sense, a solution is obtained with minimal energy (due to the ℓ_2-norm objective function), that is accurate close to the measurement area, where the data are fitted. At high frequencies, many coefficients are needed to fit these data with minimum energy. However, as we move far from the measurement area, this representation of the sound field



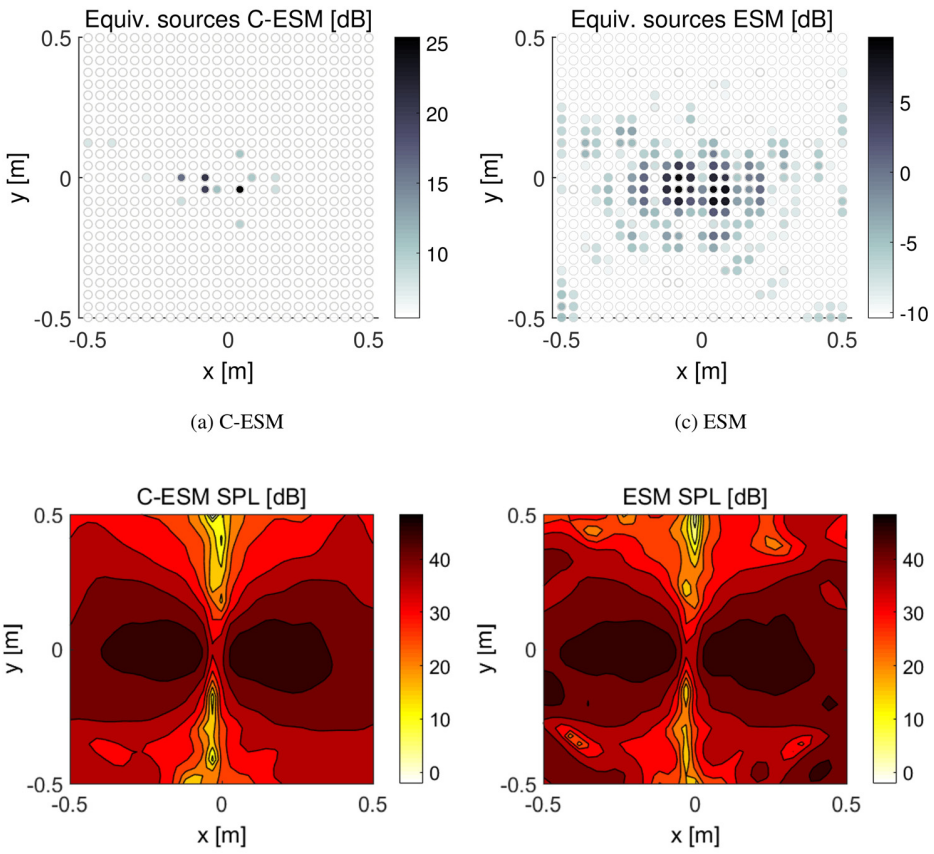
(a) C-ESM

(c) ESM

(b) C-ESM

(d) ESM

FIG. 12. (Color online) Experimental results for the dipole source at 250 Hz. (top)Estimated magnitude of the equivalent sources, and (bottom) the resulting reconstructed sound pressure, (a-b) for the proposed C-ESM methodology, and (c-d) the classical ESM solution.



(a) C-ESM (c) ESM

(b) C-ESM (d) ESM

FIG. 13. (Color online) As Fig. 12, but at 1200 Hz.

is inaccurate, and the error increases drastically. Contrarily, there are no sidelobes the CS solution of C-ESM, because of seeking a sparse vector (due to the ℓ_1-norm objective function). Consequently, the solution is accurate over a wide frequency range, beyond the conventional notions of aliasing limits/errors.[14] The solution obtained by CS (i.e., few coefficients), promotes a meaningful representation of the actual source and of the sound field, regardless of the frequency range.

V. EXPERIMENTAL RESULTS

An experiment was conducted to examine the proposed C-ESM method. Two sources of fundamentally different nature are considered, a dipole source and a classical guitar. The array used for the measurements is a Brüel & Kjær (Naerum, Denmark) Combo array of 1 m diameter, consisting of 60 $\frac{1}{4}$ in. electret microphones with pseudo-random spatial sampling.[25] The experiments were conducted in the Technical University of Denmark large anechoic chamber (Kongens Lyngby, Denmark) in February 2016.

A. Dipole source

A dipole source is examined in this section (see Fig. 11). The source consists of two loudspeaker drivers (5 in. diameter, 6.35 cm radius) mounted against each other and driven in antiphase. This source was examined in Ref. 31, and it is of interest due to it being spatially confined and highly reactive. The array was placed at $z = 7$ cm (the edge of the units is at $z = 0$ and therefore the physical center of the source at $(x, y, z) = (0, 0, -6.35)$ cm. The plane of the loudspeaker drivers oriented normal to the array plane, such that the zero pressure plane traverses the center of the array. The grid of equivalent sources used is of 25×25, and they are placed at $z = -5$ cm. The source was driven with white noise, and 1 s Hanning windows were used for the analysis. The noise floor estimate in Eqs. (11) and (12) is selected based on the highest measured noise-floor of the transducers. The reconstruction of

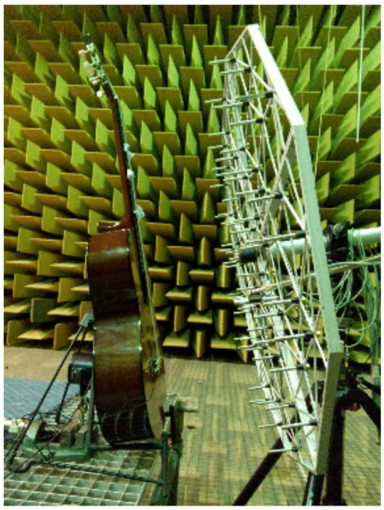


FIG. 14. (Color online) The experimental set-up for the measurements on the guitar.

the sound pressure field takes place at $z_s = 3$ cm, (i.e., near the edge of the loudspeaker units).

Figure 12 shows the estimated equivalent sources and reconstructed sound pressures based on the C-ESM and conventional ESM methods at 250 Hz. In agreement with the numerical results, it can be seen that the reconstructed pressure by the two methods is fairly similar, although the recovered equivalent sources are drastically different. In this case (contrary to the numerical results) the finite extent of the sources, as well as scattering from the setup is noticeable in the recovered equivalent sources. The C-ESM methodology estimates a fairly sparse solution, consisting on a few non-zero sources, localized at the two poles of the dipole source and near the tripod. The ESM solution consists of a similar distribution, but where the characteristic lobes and sidelobes are noticeable. As for the reconstructed sound field, the C-ESM reconstruction results in a field that is closer to the radiation pattern of a dipole.

Figure 13 shows the reconstructed sound pressures and estimated equivalent sources at a frequency of 1200 Hz. The effects of discretization start to be critical in the ESM (least-squares) estimate, and strong sidelobes appear, as can be appreciated in Fig. 13(c). Also the reconstructed pressure field by the ESM method shows the appearance of error and spatial artifacts (rapid, single point spatial variations are observable far from the source, more than a wavelength away). Conversely, the solution based on the C-ESM methodology identifies the two drivers, represented by a few dominant sources. The C-ESM solution conforms more to the radiation characteristics of the dipole source. It is noted

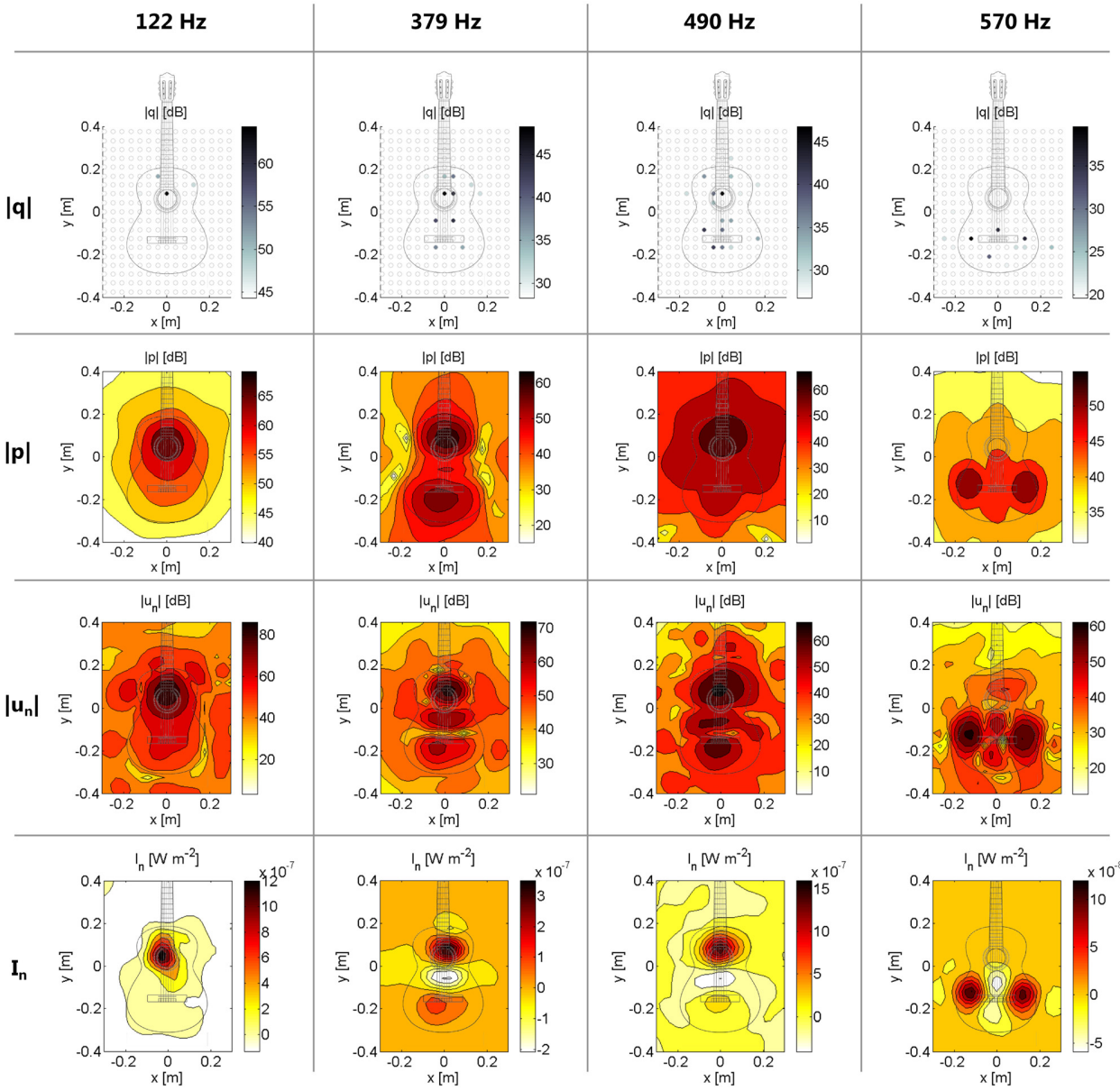


FIG. 15. (Color online) Sound field reconstruction of the field radiated by a Yamaha C45 classical guitar. The figure shows the estimated equivalent sources (top row), and reconstruction at 3 cm of the sound pressure level (mid-top row), normal particle velocity (mid-bottom row) and normal component of the active sound intensity (bottom row). Four natural frequencies are shown: the breathing mode at 122 Hz (left column), the A1 resonance at 379 Hz (mid-left column), mode in the top-plate (0,2) at 490 Hz (mid-right column), and a natural frequency at 570 Hz (right column).

that the C-ESM solution at 1200 Hz is more precise than at 250 Hz, because the observed data are less redundant at high frequencies, which favors a more accurate representation, in good agreement with the analysis in Sec. III A. In fact the Gramian of the sensing matrix $\mathbf{G}$ for the 250 Hz and 1200 Hz cases are the ones shown in Sec. III A (Fig. 3). The high coherence at 250 Hz, results in the spurious sources of high magnitude observable in Fig. 12. At 1200 Hz, the mutual coherence of the sensing matrix is lower, and the spurious sources are reduced, indicating a greater robustness to measurement noise.

B. Classical guitar

A classical acoustic guitar, Yamaha C-45 (spruce top-plate, mahogany back-plate and rosewood fingerboard) was used as a test source. The measurements took place in the large anechoic chamber at DTU (see Fig. 14). A shaker driven with white noise was placed on the back-plate of the guitar (the back-plate was chosen instead of the bridge or the top-plate, to prevent scattering). The guitar was placed on a stand on rubber pads, and the strings were muted with Basotec (open cell foam made from melanine resin). The origin of coordinates was set on the acoustic port of the guitar, on its point closest to the bridge (see Fig. 15). The measurements were conducted using the same 60 channel microphone array as in Sec. V A, and the noise-floor estimated in the same way. The array was positioned in a plane 6 cm in front of the guitar, and the reconstruction took place 3 cm in front of the top-plate. The equivalent sources were distributed on a uniform grid of 25×25, spaced in 4 cm intervals (covering an area of approximately 1 m^2), and retracted 4 cm behind the guitar's top-plate.

The classical guitar radiates very differently across frequency (Fig. 15). At low frequency, most of the radiation occurs from the acoustic port, whereas at high frequencies, it is the body of the guitar that radiates most effectively. The estimated equivalent sources, and the reconstructed pressure, velocity and intensity fields are shown for the natural frequencies 122, 379, 490, and 570 Hz. At 122 Hz, Fig. 15 shows the breathing mode of the guitar, where the sound is radiated through the acoustic port. The method (C-ESM) recovers a sparse monopole-like behavior, and the reconstructed sound pressure, velocity and intensity conform essentially to a spherical wave, as expected. At 379 Hz the A1 mode of the guitar is found, corresponding to the first acoustic resonance in the acoustic cavity (when the length of the cavity is approximately half the acoustic wavelength), and the air motion on the guitar's port is in anti-phase with the upper half of the plate.[37] This explains the strong circulation of acoustic energy as seen from the sound intensity map between the top-plate and acoustic cavity. The radiation characteristics are reminiscent of a longitudinal quadrupole, and the C-ESM methodology recovers such a source distribution. At 490 Hz we identify a mode of similar characteristics, although in this case it is due to a (0, 2) structural mode on the top-plate, and a similar interaction with the acoustic port as in the A1 mode is found.[37] Finally, at 570 Hz most of the acoustic output is radiated by the guitar's body, due to its vibration, rather than from the acoustic port. The mode at 570 Hz presumably corresponds to the fourth mode of the guitar's top-plate, due to the identified deflection shape and the resulting sound field, in agreement with Ref. 38. However, the equivalent source distribution ($\hat{\mathbf{q}}_{\ell_1}$) is not representative of the spatial extent of the acoustic source. The C-ESM solution, rather approximates the guitar's deflection shape by a set of point sources that mimic the observed sound pressure on the array. It is worth noting that this solution is not representative of the spatial extent of the source, although the reconstructed sound field conforms to the expected radiation (for example, see Fig. 9.16 of Ref. 37). This agrees with numerical results.

VI. CONCLUSIONS

A sound field reconstruction method is examined, based on the principle of wave superposition (i.e., equivalent source method) formulated as a sparse problem that can be solved via CS. The importance of having a non-redundant representation of the observed data is emphasized.

The numerical and experimental results indicate that the method can model accurately the radiation of spatially localized sources, resulting on a lower error over a greater reconstruction volume than the solutions based on conventional least-squares. The method is valid for wide band reconstruction, as the sparsity constraint suppresses sidelobes. For spatially extended sources that exhibit a pronounced modal behavior (e.g., a classical guitar at high frequencies, Sec. V B), the sparsity assumption results in a non-physical equivalent source distribution. Nonetheless, the reconstructed field still conforms to the actual wave field.

ACKNOWLEDGMENTS

This work was supported by the Danish Council for Independent Research (DFF), under the Grant No. FTP/12-126364. The authors would like to thank B&K for lending the microphone array and Jørgen Hald for discussions.

[1]E. G. Williams and J. D. Maynard, "Holographic imaging without the wavelength resolution limit," Phys. Rev. Lett. **45**, 554–557 (1980).
[2]J. 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**, 1395–1413 (1985).
[3]E. G. Williams, "Continuation of acoustic near-fields," J. Acoust. Soc. Am. **113**, 1273–1281 (2003).
[4]J. Hald, "Basic theory and properties of statistically optimized near-field acoustical holography," J. Acoust. Soc. Am. **125**, 2105–2120 (2009).
[5]S. F. Wu and J. Yu, "Reconstructing interior acoustic pressure fields via Helmholtz equation least-squares method," J. Acoust. Soc. Am. **104**, 2054–2060 (1998).
[6]A. Sarkissian, "Method of superposition applied to patch near-field acoustic holography," J. Acoust. Soc. Am. **118**, 671–678 (2005).
[7]R. Bai, "Application of BEM (boundary element method)-based acoustic holography to radiation analysis of sound sources with arbitrarily shaped geometries," J. Acoust. Soc. Am. **92**, 533–549 (1992).
[8]B. K. Kim and J. G. Ih, "On the reconstruction of the vibro–acoustic field over the surface enclosing an interior space using the boundary element method," J. Acoust. Soc. Am. **100**, 3003–3016 (1996).
[9]J. Antoni, "A Bayesian approach to sound source reconstruction: Optimal basis, regularization, and focusing," J. Acoust. Soc. Am. **131**, 2873–2890 (2012).
[10]M. Aucejo, N. Totaro, and J.-L. Guyader, "Identification of source velocities on 3d structures in non-anechoic environments: Theoretical

background and experimental validation of the inverse patch transfer functions method," J. Sound Vib. **329**, 3691–3708 (2010).
[11]I. N. Vekua, "On the completeness of the system of metaharmonic functions (in Russian)," Dokl. Akad. Nauk SSSR **90**(5), 715–718 (1953).
[12]G. H. Koopmann, L. Song, and J. B. Fahnline, "A method for computing acoustic fields based on the principle of wave superposition," J. Acoust. Soc. Am. **86**, 2433–2438 (1989).
[13]N. 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**, 3694–3705 (2006).
[14]E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," IEEE Signal Proc. Mag. **25**, 21–30 (2008).
[15]M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing* (Springer, New York, 2010), Ch. 1.
[16]D. Malioutov, M. Çetin, and A. C. Smith, "A sparse signal reconstruction perspective for source localization with sensor arrays," IEEE Trans. Signal Proc. **53**, 3010–3022 (2005).
[17]G. F. Edelmann and C. F. Gaumond, "Beamforming using compressive sensing," J. Acoust. Soc. Am. **130**, EL232–EL237 (2011).
[18]P. Simard and J. Antoni, "Acoustic source identification: Experimenting the ℓ_1 minimization approach," Applied Acoustics **74**, 974–986 (2013).
[19]A. Xenaki, P. Gerstoft, and K. Mosegaard, "Compressive beamforming," J. Acoust. Soc. Am **136**, 260–271 (2014).
[20]P. Gerstoft, A. Xenaki, and C. F. Mecklenbrauker, "Multiple and single snapshot compressive beamforming," J. Acoust. Soc. Am. **138**, 2003–2014 (2015).
[21]E. Fernandez-Grande and A. Xenaki, "Compressive sensing with a spherical microphone array," J. Acoust. Soc. Am. **139**, EL45–EL49 (2016).
[22]F. Ning, J. Wei, L. Qiu, H. Shi, and X. Li, "Three-dimensional acoustic imaging with planar microphone arrays and compressive sensing," J. Sound Vib. **380**, 112–128 (2016).
[23]A. Xenaki, E. Fernandez-Grande, and P. Gerstoft, "Block-sparse beamforming for spatially extended sources in a Bayesian formulation," J. Acoust. Soc. Am. **140**, 1828–1838 (2016).
[24]G. 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**, 1521–1534 (2012).
[25]J. Hald, "Fast wideband acoustical holography," J. Acoust. Soc. Am. **139**, 1508–1517 (2016).
[26]E. Fernandez-Grande and A. Xenaki, "The equivalent source method as a sparse signal reconstruction," in *Proceedings of Inter-Noise 2015*, San Francisco, CA (2015).
[27]N. M. Abusag and D. J. Chappell, "On sparse reconstructions in near-field acoustic holography using the method of superposition," J. Comput. Acoust. **24**, 1650009 (2016).
[28]F. P. Mechel, *Formulas of Acoustics*, 2nd ed. (Springer, New York, 2008), pp. 1040–1059.
[29]S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing* (Springer, New York, 2013), Ch. 1–6.
[30]O. Lylloff, E. Fernandez-Grande, F. Agerkvist, J. Hald, E. Tiana-Roig, and M. S. Andersen, "Improving the efficiency of deconvolution algorithms for sound source localization," J. Acoust. Soc. Am. **138**, 172–180 (2015).
[31]F. Jacobsen and H.-E. de Bree, "A comparison of two different sound intensity measurement principles," J. Acoust. Soc. Am. **118**, 1510–1517 (2005).
[32]M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.0 beta," http://cvxr.com/cvx (Last viewed 20 Dec. 2016).
[33]J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," Proc. IEEE **98**(5), 948–958 (2010).
[34]A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM J. Imag. Sci. **1**(2), 183–202 (2009).
[35]H. F. Alqadah, N. Valdivia, and E. G. Williams, "A super-resolving near-field electromagnetic holographic method," IEEE Trans. Antenn. Propag. **62**(7), 3679–3692 (2014).
[36]P. Gerstoft, F. Mecklenbräuker, A. Xenaki, and S. Nannuru, "Multi-snapshot sparse Bayesian Learning for DOA," IEEE Signal Processing Letters **23**, 1469–1473 (2016).
[37]N. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, 2nd ed. (Springer-Verlag, New York, 1998), Ch. 9.
[38]E. V. Jansson, "A study of acoustic and hologram interferometric measurement on the top plate vibrations of a guitar," Acustica **25**, 95–100 (1971).