Fast wideband acoustical holography

Jørgen Hald[a)]

Brüel and Kjær Sound and Vibration Measurement A/S, Skodsborgvej 307, DK-2850 Nærum, Denmark

Patch near-field acoustical holography methods like statistically optimized near-field acoustical holography and equivalent source method are limited to relatively low frequencies, where the average array-element spacing is less than half of the acoustic wavelength, while beamforming provides useful resolution only at medium-to-high frequencies. With adequate array design, both methods can be used with the same array. But for holography to provide good low-frequency resolution, a small measurement distance is needed, whereas beamforming requires a larger distance to limit sidelobe issues. The wideband holography method of the present paper was developed to overcome that practical conflict. Only a single measurement is needed at a relatively short distance and a single result is obtained covering the full frequency range. The method uses the principles of compressed sensing: A sparse sound field representation is assumed with a chosen set of basis functions, a measurement is taken with an irregular array, and the inverse problem is solved with a method that enforces sparsity in the coefficient vector. Instead of using regularization based on the 1-norm of the coefficient vector, an iterative solution procedure is used that promotes sparsity. The iterative method is shown to provide very similar results in most cases and to be computationally much more efficient. © *2016 Acoustical Society of America.*
[http://dx.doi.org/10.1121/1.4944757]

[JM]   Pages: 1508–1517

I. INTRODUCTION

Near-field acoustical holography (NAH) is based on performing two-dimensional (2D) spatial discrete Fourier transforms (DFT) and, therefore, the method requires a regular mesh of measurement positions. To avoid spatial aliasing problems, the mesh spacing must be somewhat less than half of the acoustic wavelength. In practice, this requirement sets a serious limitation on the upper frequency limit.

Some patch NAH methods, for example, the equivalent source method (ESM)[1] and statistically optimized near-field acoustical holography (SONAH),[2,3] can work with irregular microphone array geometries, but still require an average array-element spacing less than half the wavelength. As described by Hald,[4] this allows the use of irregular arrays that are actually designed for use with beamforming. Typically, good performance with beamforming can be achieved up to frequencies where the average array inter-element spacing is 2–3 wavelengths. A practical issue with such a solution is the fact that the patch NAH method requires measurement at a small distance to provide good resolution at low frequencies, while beamforming requires a medium-to-long distance to keep sidelobes at low levels. So for optimal wideband performance, two measurements must be taken at different distances, and separate types of processing must be used with the two measurements, making it difficult to combine the results into a single result covering the combined frequency range.

The rather new compressed sensing (CS) methods have started making it possible to use irregular array geometries

for holography up to frequencies where the average array inter-element spacing is significantly larger than half of the wavelength; see, for example, Refs. 5–7. In general, CS techniques allow reconstruction of a signal from sparse irregular samples under the condition that the signal can be (approximately) represented by a sparse subset of expansion functions in some domain, i.e., with the expansion coefficients of most functions equal to zero. Typically, the number of available samples is much smaller than the number of expansion functions, so the problem of determining the expansion coefficients is severely underdetermined. To obtain the correct solution, the solution method must enforce sparsity in the coefficient vector. This is typically done by somehow minimizing the 1-norm of the coefficient vector. In connection with acoustical holography, a set of elementary waves or elementary sources is used to represent the sound field in a given three-dimensional (3D) region, and measurements are taken with an irregular array. Chardon *et al.*[5] used a set of plane wave functions to represent the vibration of a planar star-shaped plate, while Hald[6,7] used a set of monopole point sources. A key problem in basis-function selection is to make sure that a sufficiently sparse coefficient vector can represent the sound field with sufficient accuracy.

The present paper describes a new method called wideband holography (WBH), which was introduced in Refs. 6 and 7, and which is covered by a pending patent.[8] The method is similar to the generalized inverse beamforming method published by Suzuki,[9] which uses a source model in terms of monopole and multipole point sources, and which minimizes a cost function including the 1-norm of the coefficient vector. This minimization is performed by a special iterative algorithm. WBH uses as source model a mesh of

a)Electronic mail: jorgen.hald@bksv.com

monopole point sources with less than half wavelength spacing, and instead of minimizing the 1-norm of the coefficient vector to enforce sparsity, a dedicated iterative solver is used that promotes sparsity in a different way. A main contribution of the present paper is a comparison of results between WBH and a method that solves an optimization problem with 1-norm minimization. The focus of the comparison is on the reconstruction of the sound intensity pattern close to a source surface. Section II outlines the basic theory. After an introduction to the applied array design in Sec. III, results of different simulated measurements are presented in Sec. IV, and Sec. V contains results from real measurements. Finally, Sec. VI contains a summary.

II. THEORY

Input data for patch holography processing are typically obtained by simultaneous acquisition with an array of M microphones, indexed by $m = 1,2,...,M$, followed by averaging of the $M \times M$ element cross-power spectral matrix between the microphones. For the subsequent description, we arbitrarily select a single high-frequency line f with associated cross-power matrix $\mathbf{G}$. An eigenvector/eigenvalue factorization is then performed of that Hermitian, positive-semi-definite matrix $\mathbf{G}$

$$\mathbf{G} = \mathbf{V}\mathbf{S}\mathbf{V}^H, \tag{1}$$

with $\mathbf{V}$ being a unitary matrix with the columns containing the eigenvectors $\mathbf{v}_\mu$, $\mu = 1,2,...,M$, and $\mathbf{S}$ a diagonal matrix with the real non-negative eigenvalues s_μ on the diagonal. Based on the factorization in Eq. (1), the principal component vectors $\mathbf{p}_\mu$ can be calculated as

$$\mathbf{p}_\mu = \sqrt{s_\mu}\mathbf{v}_\mu. \tag{2}$$

Just like ESM and SONAH, the WBH algorithm is applied independently to each of these principal components and, subsequently, the output is added on a power basis, since the components represent incoherent parts of the sound field. So for the subsequent description, we consider a single principal component, and we skip the index μ, meaning that input data are a single vector $\mathbf{p}$ with measured complex sound pressure values.

WBH uses a source model in terms of a set of elementary sources with associated elementary wave functions (the dictionary in CS terminology) and solves an inverse problem to identify the complex amplitudes of all elementary sources. The source model then applies for 3D reconstruction of the sound field. Here, we will consider only the case where the source model is a mesh of monopole point sources retracted to be behind/inside the real/specified source surface, i.e., similar to the model applied in ESM.[1] Figure 1(a) shows a typical setup for patch ESM, where a planar array covers a patch of the source surface, and a monopole source model inside the source surface extends somewhat beyond the measurement area. Figure 1(b) shows an alternative configuration of the equivalent sources that avoids the need for a 3D source surface geometry: A planar mesh of monopoles parallel with the array is retracted some small distance inside the
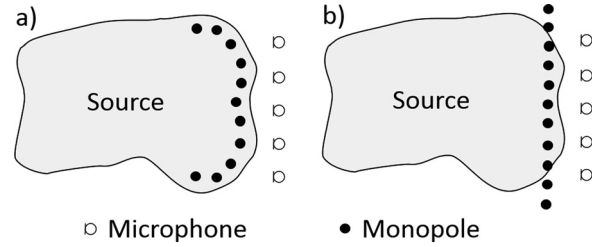

FIG. 1. Illustration of two alternative setups for patch ESM.

real source surface and covers an area slightly larger than the array.

With A_{mi} representing the sound pressure at microphone m due to a unit excitation of monopole number i, the requirement that the modeled sound pressure at microphone m must equal the measured pressure p_m can be written as

$$p_m = \sum_{i=1}^{I} A_{mi}q_i. \tag{3}$$

Here, I is the number of point sources, and q_i, $i = 1,2,...,I$, are the complex amplitudes of these sources. Equation (3) can be rewritten in matrix-vector notation as

$$\mathbf{p} = \mathbf{A}\mathbf{q}, \tag{4}$$

where $\mathbf{A}$ is an $M \times I$ matrix containing the quantities A_{mi}, and $\mathbf{q}$ is a vector with elements q_i. In CS terminology, the matrix $\mathbf{A}$ is called the sensing matrix.

When doing standard patch holography calculations using ESM, Tikhonov regularization is typically applied to stabilize the minimization of the residual vector $\mathbf{p}$-$\mathbf{A}\mathbf{q}$. This is done by adding a penalty proportional to the 2-norm of the solution vector when minimizing the residual norm

$$\underset{\mathbf{q}}{\text{Minimize}} \quad \|\mathbf{p} - \mathbf{A}\mathbf{q}\|_2^2 + \theta^2\|\mathbf{q}\|_2^2. \tag{5}$$

A very important property of that problem is the fact that it has the simple analytic solution

$$\mathbf{q} = [\mathbf{A}^H\mathbf{A} + \theta^2\mathbf{I}]^{-1}\mathbf{A}^H\mathbf{p}, \tag{6}$$

where $\mathbf{I}$ is a unit diagonal matrix, and H represents Hermitian transpose. A suitable value of the regularization parameter θ for given input data $\mathbf{p}$ can be identified automatically, for example, by use of generalized cross validation (GCV); see Gomes and Hansen.[10] When using a specific irregular array well above the frequency of half wavelength average microphone spacing, the system of linear equations in Eq. (4) is, in general, strongly underdetermined, because the monopole mesh must have spacing less than half of the wavelength, i.e., much finer than the microphone grid. During the minimization in Eq. (5), the undetermined degrees of freedom will be used to minimize the 2-norm of the solution vector. The consequence is a reconstructed sound field that matches the measured pressure values at the microphone positions, but with minimum sound pressure elsewhere. Estimates of, for example, sound power will

J. Acoust. Soc. Am. **139** (4), April 2016

Jørgen Hald   1509

therefore be much too low. Another effect is ghost sources that help focus radiation toward the microphones. This will be illustrated in Sec. IV A by simulated measurements.

If the true source distribution is sparse (with a majority of elements in $\mathbf{q}$ equal to zero), or close to sparse, the above phenomena can be alleviated by replacing the 2-norm in the penalty term of Eq. (5) by a 1-norm

$$\underset{\mathbf{q}}{\text{Minimize}} \quad \|\mathbf{p} - \mathbf{Aq}\|_2^2 + \theta^2 \|\mathbf{q}\|_1; \tag{7}$$

see, for example, Refs. 5 and 9. The minimization of the 1-norm of the solution vector will have the effect of favoring sparse solutions. Important problems related to this formulation are the lack of an analytic solution and the fact that no good tool is available to identify an optimal value of the regularization parameter θ for given input data $\mathbf{p}$. An equivalent problem was solved by Chardon *et al.*[5]

$$\underset{\mathbf{q}}{\text{Minimize}} \quad \|\mathbf{q}\|_1 \text{ subject to } \|\mathbf{p} - \mathbf{Aq}\|_2 \leq \delta. \tag{8}$$

Here, however, the parameter δ is difficult to determine. In cases where the applied source model cannot represent the full measured sound field (for example, due to sources outside the measurement region or reflections), a rather large value of δ may be needed in order that the constraint $\|\mathbf{p} - \mathbf{Aq}\|_2 \leq \delta$ can be fulfilled. Instead of requiring a small 2-norm of the residual vector

$$\mathbf{r}(\mathbf{q}) \equiv \mathbf{p} - \mathbf{Aq}, \tag{9}$$

we can alternatively require a solution close to a minimum of the squared residual function F

$$F(\mathbf{q}) \equiv \frac{1}{2} \|\mathbf{r}(\mathbf{q})\|_2^2 = \frac{1}{2} \|\mathbf{p} - \mathbf{Aq}\|_2^2. \tag{10}$$

Such a minimum will be characterized by the gradient vector $\nabla F(\mathbf{q})$

$$\nabla F(\mathbf{q}) = -\mathbf{A}^H(\mathbf{p} - \mathbf{Aq}), \tag{11}$$

having a small norm. Thus, instead of Eq. (8), we solve the problem

$$\underset{\mathbf{q}}{\text{Minimize}} \quad \|\mathbf{q}\|_1 \text{ subject to }$$
$$\|\nabla F(\mathbf{q})\|_2 = \|\mathbf{A}^H(\mathbf{p} - \mathbf{Aq})\|_2 \leq \delta. \tag{12}$$

The optimization problem of Eq. (12) is convex and can be solved by available MATLAB libraries. In the present paper, the CVX library has been used (see Refs. 5 and 11) so the method will just be called "CVX." The computational demand is, however, significantly higher than for the Tikhonov problem in Eq. (5) because no analytic solution exists. According to experience, a good way of defining the parameter δ in Eq. (12) is

$$\delta = \varepsilon \|\nabla F(\mathbf{q} = 0)\|_2 = \varepsilon \|\mathbf{A}^H \mathbf{p}\|_2, \tag{13}$$

where ε is a small number. Its value must be chosen small enough to guarantee that we get close to a minimum, but

large enough that we do not enforce a modeling of measurement noise/errors. The requirement on the gradient norm in Eq. (12) occurs also in the stopping criterion of the iterative solution method implemented in WBH [see Eq. (14) below], but only as one of several conditions that will imply stopping.

A main idea behind the iterative WBH algorithm is to remove/suppress the ghost sources associated with the real sources in an iterative solution process, starting with the strongest real sources. WBH applies a steepest descent iteration to minimize the quadratic residual function $F(\mathbf{q})$; see Appendix A for details. In the first step, a number of real as well as ghost sources will appear in $\mathbf{q}$. When using irregular array geometries, the ghost sources will in general be weaker than the strongest real source(s). We can therefore suppress the ghost sources by setting all components in $\mathbf{q}$ below a certain threshold to zero. The threshold is computed as being a number of decibels D below the amplitude of the largest element in $\mathbf{q}$. Initially, D is set to a very small value D_0, and it is then increased during the iteration, typically by the same number of decibels ΔD in every step. Application of the algorithm to many different types of sources and the use of several different arrays over a broad range of frequencies has shown that $D_0 = 0.1$ and $\Delta D = 1$ will work very well in most cases. Notice that the dynamic range limitation will gradually disappear during the iteration, so it has the role of steering the iteration toward a sparse solution in its initial phase. Therefore, the solution provided will, in most cases, not be strictly sparse.

Figure 2 contains a block flowchart diagram for the iteration, which is stopped when

$$D > D_{\text{max}} \quad \text{or} \quad \|\nabla F(\mathbf{q})\|_2 \leq \varepsilon \|\nabla F(\mathbf{0})\|_2 = \varepsilon \|\mathbf{A}^H \mathbf{p}\|_2, \tag{14}$$

where D_{max} is an upper limit on D and ε is a small number. The following values have been found to work, in general,
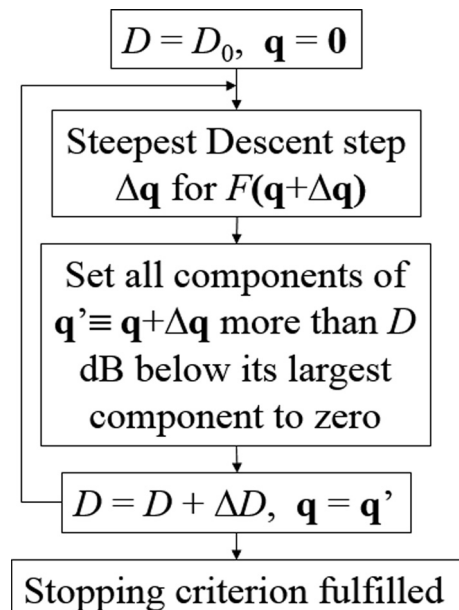

FIG. 2. Flowchart of the modified steepest descent algorithm used in WBH.

very well: $D_{max} = 60$ and $\varepsilon = 0.01$. Typically, the steepest descent method has very slow final convergence because of zigzagging and, therefore, the first of the two criteria in Eq. (14) will usually be fulfilled first. D_{max} can be changed to match the quality of data, but the choice does not seem to be critical. $D_{max} = 60$ has been found to support the identification of weak sources, even when measurements are slightly noisy. Larger values do not seem to improve much. Smaller values may be required for very noisy data.

Starting with only 0.1 dB dynamic range means that only the very strongest source(s) will be retained, while all related ghost sources will be removed. When we use the dynamic range limited source vector as the starting point for the next iteration, the components of the residual vector related to the very strongest source(s) have been reduced and, therefore, the related ghost sources have been reduced correspondingly. Increasing the dynamic range will then cause the next level of real sources to be included, while suppressing the related ghost sources, etc.

After the termination of the above algorithm based on steepest descent directions, a good estimate of the basic source distribution has been achieved. The typical zigzagging progress in the last steepest descent steps means that often the direction given as the sum of the last two steps will support good progress. The step-size that minimizes the quadratic function $F(\mathbf{q})$ along that direction can be calculated analytically, and that so-called extrapolation step is used as part of WBH in the present work.

A few conjugate gradient iteration steps without dynamic range limitation can then optionally be performed to ensure convergence to a point very close to a minimum of $F(\mathbf{q})$. See Appendix B for details on the algorithm. Usually, the effect on the source model and the modeled sound field is relatively small, because the primary steepest descent algorithm has already reduced the residual to be close to a minimum. However, it ensures that full convergence has been achieved. The stopping criteria used with the conjugate gradient iteration are

$$
\begin{aligned}
&\|\nabla F(\mathbf{q} + \Delta\mathbf{q})\|_2 < \varepsilon \|\mathbf{A}^H \mathbf{p}\|_2 \quad \text{or} \\
&\|\nabla F(\mathbf{q} + \Delta\mathbf{q})\|_2 > \|\nabla F(\mathbf{q})\|_2,
\end{aligned}
\tag{15}
$$

where $\Delta\mathbf{q}$ is the latest step. In comparison with the CVX method defined in Eqs. (12) and (13), notice that selection of too small a value of ε will not prevent termination of the conjugate gradient algorithm of WBH. In that case, it will stop when the gradient norm starts increasing. When that happens, the last step is discarded. Notice that the conjugate gradient steps have no dynamic range limitation and therefore do not retain sparsity. This will compensate for the fact that, in most cases, the real source distribution is not sparse. In the context of the present paper, conjugate gradient iterations are always used and they are considered as part of the WBH algorithm.

Both the WBH and the CVX method will have problems separating two compact, closely spaced sources at low frequencies when only a limited dynamic range of data can be used, which sets a lower limitation on the parameter ε. In

that case, the combined sound pressure distribution from the two real sources across the array cannot be distinguished from the pressure distribution from a single source at an intermediate position. With a sufficiently large applicable dynamic range of data, it would have been possible. As a consequence, the two sources will be more or less represented by a single source at the intermediate position. An example of this is given in Sec. IV B. The phenomenon is worse in the case of WBH than for CVX for the following reason: It is quite easy to show that the first steepest descent direction $\Delta\mathbf{q}$ (see Fig. 2) is equivalent to a delay and sum (DAS) beamforming map, which has poor low-frequency resolution. The first source identified by WBH will therefore be located at the peak of a DAS map, which, in this case, is somewhere between the two real sources. At higher frequencies, DAS has good resolution, so therefore the problem is of much less practical importance.

The fact that the WBH algorithm will often replace two (or more) physical low-frequency sources by a single central source might lead the user to draw wrong conclusions about the root cause of the noise. Use of the traditional Tikhonov regularization of Eq. (6), i.e., a standard ESM algorithm, will, in that case, typically show a single large source area covering the two (or more) physical sources. To minimize the risk of misleading results, it is recommended to use the standard ESM solution up to a transition frequency at approximately 0.7 times the frequency of half the wavelength given by the average array inter-element spacing (i.e., spacing $\approx 0.35\lambda$), and above that transition frequency switch to the use of WBH. This is illustrated by the simulated measurement in Sec. IV B. The fact that the CVX method exhibits similar behavior, although to a smaller extent, means that it should probably also be supplemented by, typically, an ESM algorithm at low frequencies. The proposed switchover frequency has been found to work well for many different types of sources and different array geometries.

III. ARRAY DESIGN

As described in the Introduction, the method of the present paper follows the principles of CS being based on measurements with a random or pseudo-random array geometry in combination with an enforced sparsity of the coefficient vector of the source model. The array geometry used in the simulated measurements of the present paper is shown in Fig. 3. It has 12 microphones uniformly distributed in each one of five identical angular sectors. The average element spacing is ~12 cm, implying a low-to-high transition frequency close to 1 kHz (where 0.35λ is close to 12 cm). The geometry has been optimized for minimum sidelobe level with DAS beamforming measurements up to 6 kHz as described in Ref. 4. This optimization guarantees a maximum ability of the array to distinguish plane waves incident from different directions. If the measurement distance is not too short, the ability to distinguish point sources in different directions will also be good.

An important finding from simulated measurements with the chosen type of array design is that the measurement
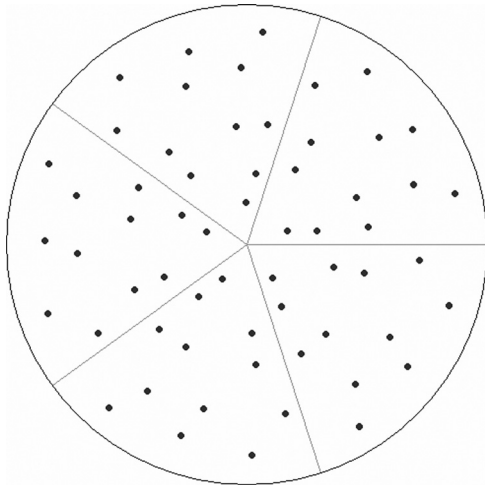
J. Acoust. Soc. Am. **139** (4), April 2016

Jørgen Hald   1511

FIG. 3. Geometry of the applied planar pseudo-random 60-element microphone array with 1 m diameter.

FIG. 4. (Color online) Setup for simulated measurement on a single monopole point source.

distance should not be shorter than approximately a factor two times the average microphone spacing for the method to work well at the highest frequencies. A factor of three is even better, and distances up to, typically, 0.7 times the array diameter work fine. When the measurement distance is increased, each source in the WBH source model will expose a greater part of the array. This means that a particular point source will create an even sound pressure over a greater part of the array, making more effective use of the resolution capabilities of the full array. In addition, a longer measurement distance will probably cause the non-sparsity of the real source to be less "visible" in the measured data, causing fewer problems in fitting a sparse model to the data. To get acceptable low-frequency resolution, however, the measurement distance should not be too long either, so, overall, the best distance seems to be two to three times the average array inter-element spacing. These considerations hold true also when the CVX algorithm is used. A thorough investigation is not given in the present paper, so this could be a topic for a future publication, together with a more extensive investigation of the sound-field reconstruction errors.

IV. SIMULATED MEASUREMENTS

All CVX and WBH calculations in the present paper were performed using $D_0 = 0.1$, $\Delta D = 1.0$, $D_{max} = 60$, and $\varepsilon = 0.01$.

A. Single monopole point source

The aim of the single-monopole simulated measurement is to demonstrate: (i) What happens if Tikhonov regularization is applied above the frequency of half wavelength average array element spacing? (ii) How much and which kind of improvement is achieved by applying the sparsity promoting CVX and WBH algorithms?

As illustrated in Fig. 4, we consider a setup with a monopole point source located on the array axis at 28 cm distance from the array plane, while the source-model mesh is at 27 cm distance, and the sound field in reconstructed in a
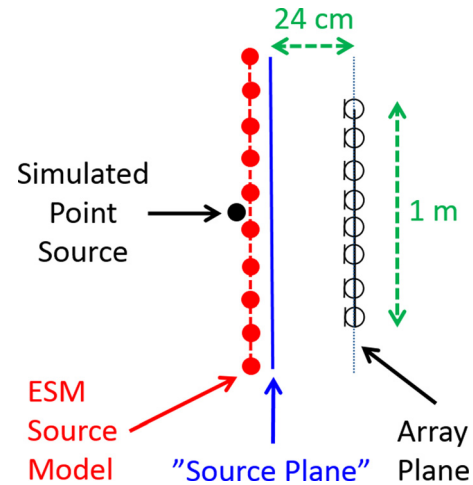
"source plane" 24 cm from the array plane. The reconstruction mesh has 51 columns and 51 rows with 2 cm spacing, covering a 1 m × 1 m area centered on the array axis. The source-model mesh is similar, i.e., with 2 cm spacing, but it is extended by six rows/columns in all four directions. In total, $63 \times 63 = 3969$ complex point-source amplitudes must be determined from the 60 measured complex sound-pressure values. No measurement errors/noise was added.

The left column of Fig. 5 shows the 4 kHz sound intensity maps on the source plane, while the right column shows the corresponding sound pressure maps on the measurement plane. From the top and downward, the true maps are shown, followed by the reconstructions using Tikhonov regularization (using 20 dB dynamic range), CVX, and at the bottom WBH. The CVX and WBH maps are both very close to the true maps, as could be expected in the present case, where the source-model plane is only 1 cm from the monopole point source. The sound intensity reconstruction based on Tikhonov regularization shows a small low-level peak at the true source position, but, in addition, there are quite a lot of ghost sources. These ghost sources are responsible for focusing the modeled sound field toward the microphones to produce the measured pressure at these positions with a minimum of radiated power. This is evident from the map of the sound pressure in the array plane.

Area integration of sound intensity maps like those in Fig. 5 leads to the sound power spectra compared in Fig. 6. Here, the spectra from the three reconstruction methods have been normalized by the spectrum from the true intensity map. CVX and WBH produce accurate sound power estimates over the entire range from 100 Hz to 6.4 kHz, while Tikhonov regularization leads to acceptable accuracy up to ∼1 kHz. Above that frequency, an increasing underestimation is observed due to the increasing ability of the source model to establish the measured pressure at the microphones and low pressure in all other directions. Calculation times for the 64 frequencies represented in Fig. 5, using MATLAB implementations of the CVX and WBH methods, were 829 s for CVX and 32 s for WBH.
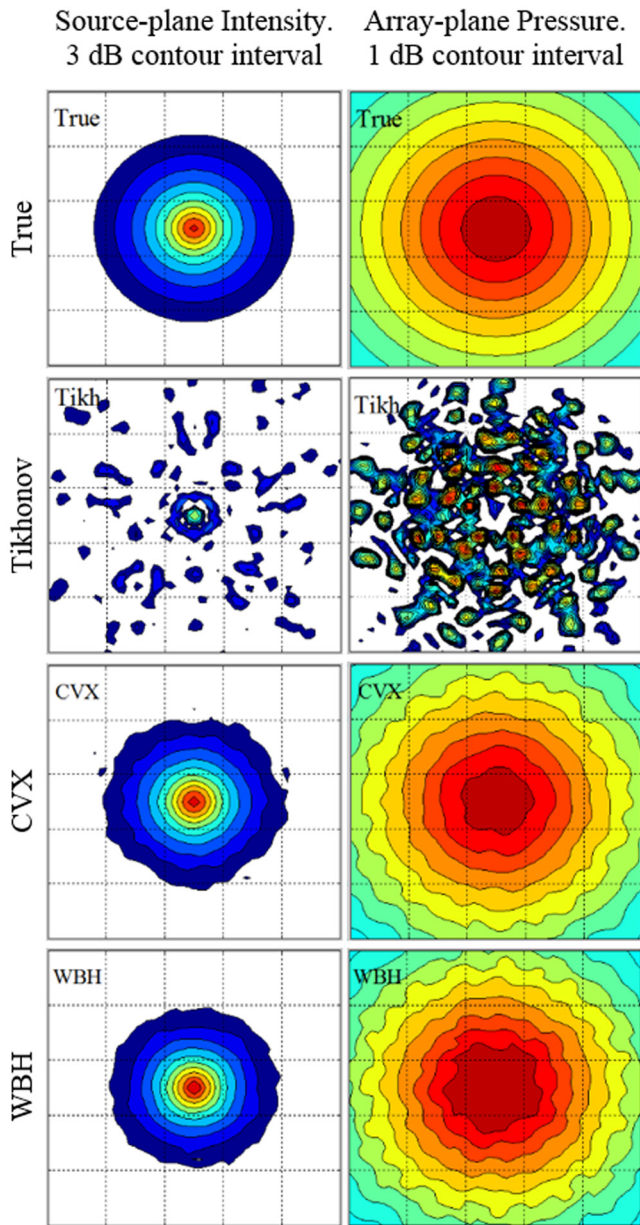
                                   Jørgen Hald

FIG. 5. (Color online) Results from a simulated measurement on a single monopole point source. All plots in a column have equal scaling.
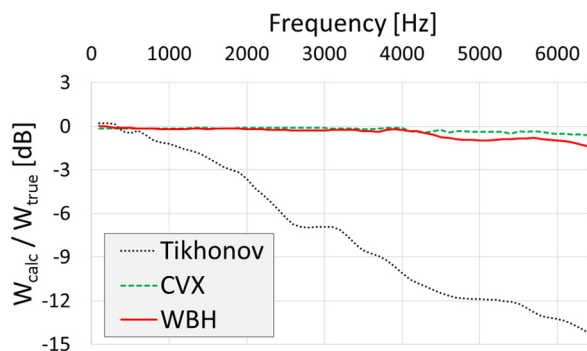

FIG. 6. (Color online) Estimated sound power spectra relative to the true sound power spectrum. All spectra were obtained by area integration of the intensity maps in Fig. 5.

B. Two coherent, in-phase monopole sources of equal level

The aim of this simulated measurement is to demonstrate some important resolution properties at low frequencies. For that purpose, we use a setup similar to that of Fig. 4, but with two sources within a mapping area of half the dimensions. The two monopole point sources are located 29 cm in front of the array plane at (x,y) coordinates $(15,15)$ cm and $(-15,-15)$ cm relative to the array axis, while the source-model mesh is at a distance of 25.5 cm, and the sound field is reconstructed in a "source plane" 24 cm from the array plane. Thus, in this case, the real sources are 3.5 cm behind the source model. The reconstruction mesh has 51 columns and 51 rows with 1 cm spacing, covering a $0.5\,\text{m} \times 0.5\,\text{m}$ area centered on the array axis, and the source-model mesh is similar, i.e., with 1 cm spacing, but it is extended by six rows/columns in all four directions. Random noise was added to the complex microphone pressure data at a level 30 dB below the average sound pressure across the microphones.

Figure 7 shows sound intensity maps at 400 Hz. The top row contains to the left the true intensity and to the right the intensity obtained using Tikhonov regularization with 20 dB dynamic range. With the applied combination of source separation, measurement distance and dynamic range, the two sources are not resolved when using Tikhonov, although it is close. The map indicates, however, the shape of the source distribution. The bottom row contains to the left the CVX reconstruction and to the right the WBH result. Both methods put a concentrated false source in the middle between the two real sources plus two sources close to the real sources. This phenomenon, which is strongest for WBH, could lead to wrong conclusions about the origin of the measured noise. So, as described at the end of Sec. II, it is
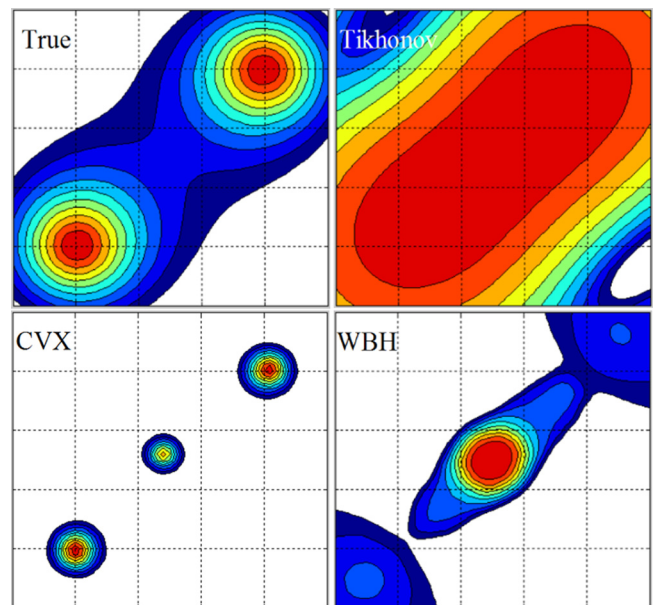

FIG. 7. (Color online) True and reconstructed sound intensity maps for two coherent, in-phase monopole sources of equal amplitude. All plots have 1.5 dB contour interval, but different thresholds. The area-integrated sound power is almost identical for all plots.

J. Acoust. Soc. Am. **139** (4), April 2016

Jørgen Hald   1513

recommended to have at least an option for use of standard ESM (with Tikhonov regularization) at the low frequencies instead of CVX or WBH.

C. Two coherent, in-phase monopole sources with 10 dB level difference

A main purpose of this section is to investigate the ability of the CVX and WBH methods to identify weak sources in the presence of strong ones. We use the same setup as in Sec. IV B, except that the lower left source (source 2) is now assigned an excitation 10 dB below that of the upper right source (source 1). Figure 8 shows the true and the reconstructed sound intensities on the "source plane" at 5 kHz with a 20 dB display range. The two sources are well identified by both CVX and WBH, and the two methods show very similar results. The maps actually look much the same at all frequencies between 1 kHz and 5 kHz. Sound power integration areas are shown with line style corresponding to the associated sound power spectra in Fig. 9. Except for the weak source 2 at the lowest frequencies, the two reconstruction methods estimate almost the same sound power spectra for the two sources. As described above, a standard ESM algorithm should be used anyway at the lowest frequencies – for the present array up to 1 kHz. The apparent small overestimation of the sound power for source 1 up to around 4 kHz is probably, instead, an underestimation of the true power: The true intensity map is smoother and therefore some of the power of source 1 will be outside the associated integration area. Apart from a 2.5 dB dip around 4 kHz in the estimated power for source 2, the accuracy is good up to around 5 kHz, and above that frequency, an increasing underestimation is observed. The maximum frequency of the present array (with 12 cm average microphone spacing) in connection with the SONAH and ESM algorithms is ~1.2 kHz, so,
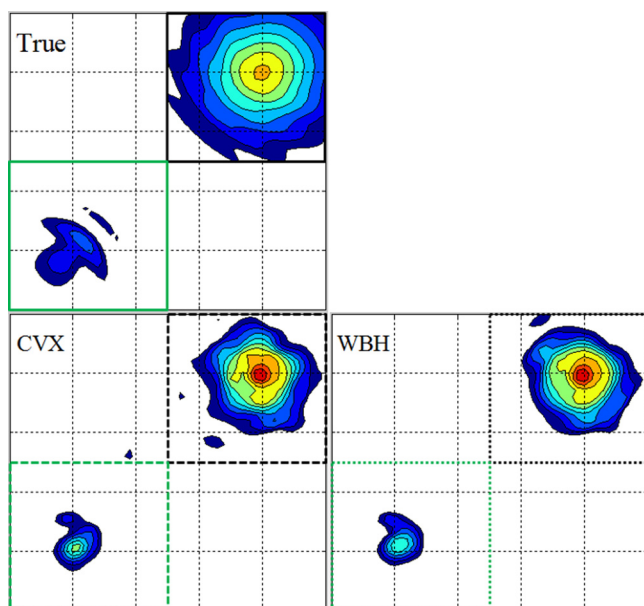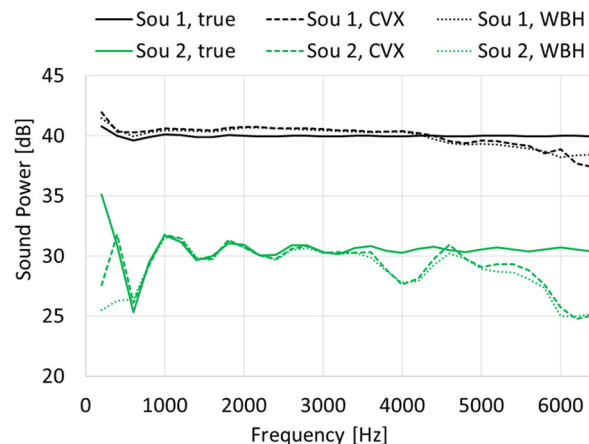
FIG. 9. (Color online) Area-integrated sound power spectra for the areas of Fig. 8.

apparently, the CVX and WBH methods extend the frequency range by a factor of around 4.

The calculation times for the 32 frequencies represented in Fig. 9 were 490 s for CVX and 16 s for WBH, so again WBH is faster by approximately a factor of 30. Another advantage of WBH is the already mentioned smaller sensitivity of WBH to the specified target reduction in the gradient norm: Too small a value of ε causes the CVX method to become unstable.

D. Plate in a baffle

The aim of the simulated plate measurements is to show that both the CVX and the WBH methods can give quite good results even when the true source distribution is not sparse. As an example of a more distributed source, a baffled, center-driven, simply supported, 6 mm thick, 40 cm × 40 cm aluminum plate has been used. The coincidence frequency for the plate is at 2026 Hz. The vibration pattern was calculated using the formulation by Williams[12] and, subsequently, the radiated sound field was obtained using the discretized Rayleigh integral, approximating the plate velocity distribution by 161 × 161 monopole point sources. This allowed the microphone sound-pressure values and the "true" pressure and particle velocity in a reconstruction plane 1 cm above the plate to be calculated. As for the simulated measurements on two monopole point sources, random noise was added to the complex microphone pressure data at a level 30 dB below the average sound pressure across the microphones. The reconstruction mesh had 41 × 41 points with 1 cm spacing, covering exactly the plate area, and the array was placed 24 cm above the plate. For the WBH sound field reconstruction, a source model comprising 53 × 53 monopole point sources with 1 cm spacing was located 1 cm below the plate.

Figure 10 shows the true sound intensity and the corresponding CVX and WBH reconstructions at 3 and 4 kHz with a 20 dB display range. Overall, the reconstruction is good, with a little too high weight on the central area, the two methods performing again very equally. At 4 kHz, the reconstructed intensity patterns start getting distorted because the complexity of the vibration pattern becomes too

FIG. 8. (Color online) True and reconstructed sound intensities in the "source plane." Display range is 20 dB with 2 dB contour interval. All three plots use the same scale. Source 1 is the stronger source in the upper right corner.
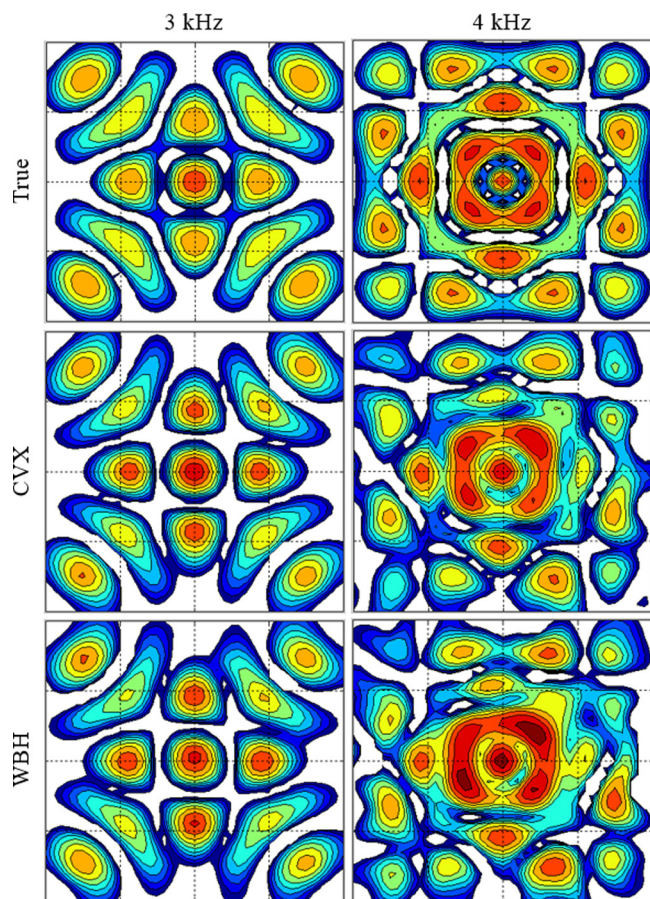
FIG. 10. (Color online) Contour plots at 3 and 4 kHz of sound intensity in the reconstruction plane 1 cm above the plate. Display range is 20 dB with 2 dB contour interval. For each frequency, the same scale is used.

high in relation to the data provided by the array. As mentioned earlier, the reconstruction accuracy at the highest frequencies can be improved by an increase of the measurement distance up to three times the array inter-element spacing, but, of course, at the expense of slightly poorer low-frequency resolution.

Figure 11 shows the relative sound power spectrum of the CVX and WBH reconstructions: At each frequency, the reconstructed and true sound intensity maps (as shown in Fig. 10) have been area integrated over the entire plate, and the ratio between the estimated and the true sound power

values have been plotted in decibels. There is a consistent small underestimation, but up to 5 kHz it remains within 2 dB. Above 5 kHz, the underestimation increases rapidly, in particular, for the CVX based algorithm. The calculation time for the 32 frequencies represented in Fig. 11 was 238 s for CVX and 9 s for WBH.

V. REAL MEASUREMENT ON TWO SMALL LOUDSPEAKERS

Figure 12 shows two Brüel and Kjær Mouth Simulators type 4227 set up 36 cm from the array and with 12 cm separation between the two sources. Thus, the measurement distance has been increased to three times the average microphone spacing. The two sources were excited from two independent stationary-random white-noise generators adjusted to equal levels. Beyond the array measurement, a scan was also performed with a two-microphone sound intensity probe across a plane 7 cm from the two sources. 13 × 6 positions with 3 cm spacing were measured, covering an area of 36 cm × 15 cm. The measurements were performed in a normal room with no acoustical treatment.

The array measurement consisted in simultaneously recording 10 s of time data with 12.8 kHz bandwidth from all array microphones. As described in Sec. II, the processing started with averaging of the 60 × 60 element cross-spectral matrix between all array microphones. Then, a principal component decomposition of that matrix was performed, and the WBH algorithm was applied to each significant component. In the present case of two independently excited sources, there were two such significant principal components. The planar WBH reconstruction mesh was in a source plane parallel with the array plane and touching the source units, and it consisted of 41 × 41 points with 1 cm spacing. The source model mesh was similar to the calculation mesh, but shifted 1.5 cm away from the array and extended by six rows/columns in all four directions.

Figure 13 shows contour plots of the reconstructed sound intensity for the two 1/3-octave bands at 4 and 5 kHz. WBH was applied to FFT spectra with 32 Hz line width, and 1/3 octaves were then synthesized. The significant level difference between the two source units in the 5 kHz band was consistent with beamforming processing of the same array data and with the intensity maps from the intensity probe scan.
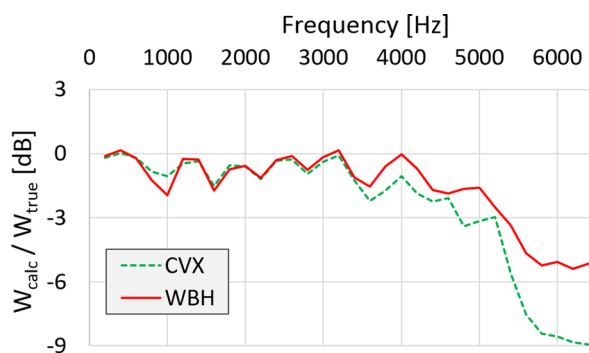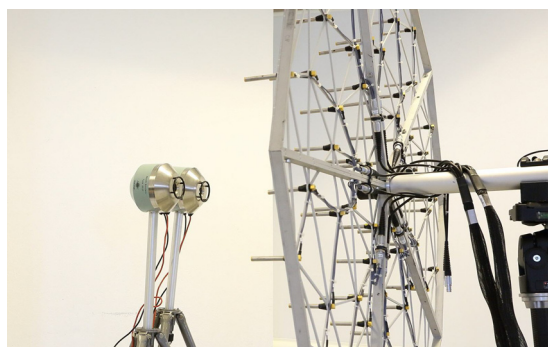

FIG. 11. (Color online) Reconstructed relative to true sound power of the plate in decibels. All power values were obtained by area integration of maps like those in Fig. 10.

FIG. 12. (Color online) Sixty-channel array at 36 cm distance from two Brüel & Kjær Mouth Simulators type 4227.

J. Acoust. Soc. Am. **139** (4), April 2016
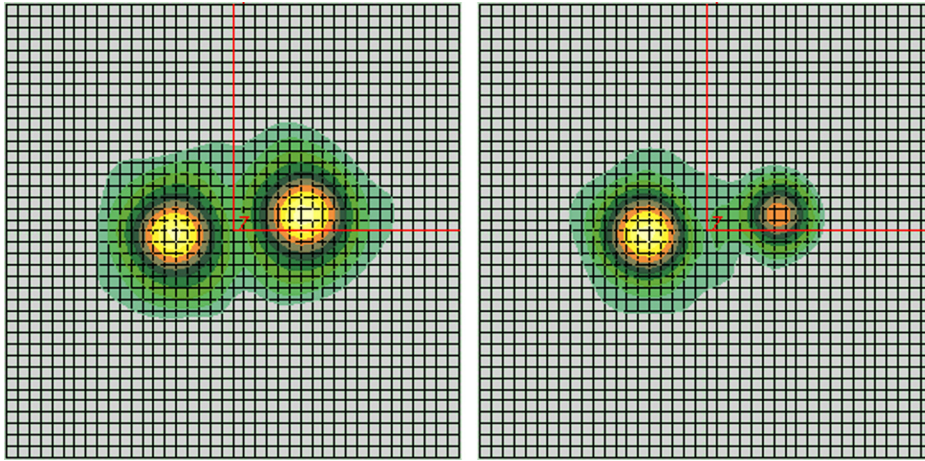
Jørgen Hald   1515

FIG. 13. (Color online) Sound intensity maps from WBH for the 1/3-octave bands 4 kHz (left) and 5 kHz (right).

Figure 14 compares the sound power spectrum from the intensity probe scan with the sound power spectrum from the WBH reconstruction. Both were obtained by area integration of sound intensity maps. However, where the WBH map covers a relatively large area in the source plane, the intensity probe map covers a rather limited area at 7 cm distance. Consequently, the WBH result will be an estimate of the total sound power radiated to a hemisphere, while the intensity-probe result will include only a part of that power. The generally slightly higher level of the WBH spectrum in Fig. 14 should therefore be expected. The part of the WBH spectrum below ∼1 kHz was obtained using standard ESM, so the iterative WBH algorithm seems to provide more accurate sound power estimates than standard ESM. As a conclusion, the sound power estimates from WBH are very accurate.

VI. SUMMARY

An iterative algorithm has been described for sparsity promoting NAH over a wide frequency range based on the use of an optimized pseudo-random array geometry. The method, which is called WBH, can be seen as an example of CS. The algorithm has been tested by a series of simulated measurements on point sources and on a plate in a baffle. Very good results were, in general, obtained at frequencies up to four times the normal upper limiting frequency for use of the particular array with holography. The focus has been on the ability to locate and quantify the main sources (source areas) in terms of sound power within approximately a 10 dB

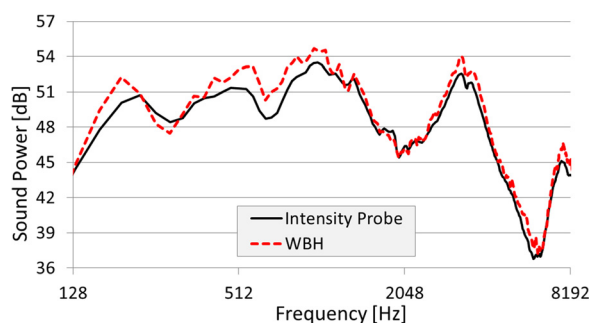FIG. 14. (Color online) Comparison of narrow-band sound power spectra from the intensity probe scan and from WBH processing of the array data.

dynamic range. The method was found to work surprisingly well with distributed sources, such as vibrating plates. Typical application areas could be engines and gearboxes, where measurements at close range are often not possible, and the method seems to work very well at the distances that are typically realistic in such applications.

The iterative WBH algorithm was shown to provide sound field reconstructions almost identical to those of a conventional CS algorithm, where an optimization problem must be solved, involving minimization of the 1-norm of the solution vector. In the present work, such optimization problems have been solved using the CVX MATLAB toolbox. For all the considered examples, the computation time of the CVX-based method were approximately 30 times longer than those of the iterative WBH method. In addition, the stopping criteria of the iterative WBH algorithm support the reconstruction of a large dynamic range without the risk of introducing numerical instability. This is not possible in the CVX-based approach, where a fixed dynamic range must be specified.

It was shown in the present paper that it may be advantageous to supplement both the CVX and the WBH algorithms with, typically, a standard ESM algorithm at the lowest frequencies in order to avoid misleading ghost sources in super-resolution results.

Engine and gearbox measurements are characterized by having sources at different distances. The sensitivity of the WBH algorithm to sources located outside the assumed source plane was therefore investigated in Ref. 6. In general, the estimation of sound power was found not to be sensitive to sources being within reasonable distances from the assumed source plane.

APPENDIX A: STEEPEST DESCENT ALGORITHM

From a current approximation $\mathbf{q}$ to the complex source-model amplitudes, we wish to find the step length s in the direction $\mathbf{w}$ of steepest descent for the squared residual function F. The steepest descent direction is just the negative gradient of F

$$\mathbf{w} \equiv -\nabla F(\mathbf{q}) = \mathbf{A}^H(\mathbf{p} - \mathbf{A}\mathbf{q}) = \mathbf{A}^H\mathbf{r}, \tag{A1}$$

Jørgen Hald

with **r** being the residual vector. In the calculation of s, we need also the vector **g** defined as

$$\mathbf{g} \equiv \mathbf{Aw}. \tag{A2}$$

Using these quantities and the definition of F in Eq. (10), we get after some manipulation

$$F(\mathbf{q} + s\mathbf{w}) = \frac{1}{2}\left[s^2\mathbf{g}^H\mathbf{g} - 2s\mathbf{w}^H\mathbf{w} + \mathbf{r}^H\mathbf{r}\right]. \tag{A3}$$

From Eq. (A3), the step length s to the minimum of F along the direction **w** can be easily found by setting the derivative of $F(\mathbf{q} + s\mathbf{w})$ with respect to s equal to zero. As a result, we get

$$s = \frac{\mathbf{w}^H\mathbf{w}}{\mathbf{g}^H\mathbf{g}}. \tag{A4}$$

So finally, the steepest descent step $\Delta\mathbf{q}$ is calculated as

$$\Delta\mathbf{q} = s\mathbf{w}. \tag{A5}$$

APPENDIX B: CONJUGATE GRADIENT ALGORITHM

There is a rich literature on such algorithms; see, for example, Refs. 13 and 14. The algorithms solve a set of real or complex linear equations, where the coefficient matrix is positive definite and real symmetric or complex Hermitian. In order to bring the system of equations in Eq. (4) into a tractable form, we multiply by $\mathbf{A}^H$ on both sides, which leads to the normal equations

$$\mathbf{A}^H\mathbf{p} = \mathbf{A}^H\mathbf{Aq}. \tag{B1}$$

Comparison with Eq. (11) shows that solution of Eq. (B1) is equivalent to finding the vector **q** for which the gradient of $F(\mathbf{q})$ equals zero

$$\nabla F(\mathbf{q}) = -\mathbf{A}^H(\mathbf{p} - \mathbf{Aq}) = \mathbf{0}. \tag{B2}$$

The system of linear equations in Eq. (B1), with starting point obtained from the modified steepest descent algorithm, can be solved by application of the conjugate gradient algorithm.

[1]A. Sarkissian, "Method of superposition applied to patch near-field acoustical holography," J. Acoust. Soc. Am. **118**, 671–678 (2005).

[2]J. Hald, "Basic theory and properties of statistically optimized near-field acoustical holography," J. Acoust. Soc. Am. **125**, 2105–2120 (2009).

[3]J. Hald, "Scaling of plane-wave functions in statistically optimized near-field acoustic holography," J. Acoust. Soc. Am. **136**, 2687–2696 (2014).

[4]J. Hald, "Array designs optimized for both low-frequency NAH and high-frequency beamforming," in *Proceedings of Inter-Noise* (2004), paper 425.

[5]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).

[6]J. Hald, "Wideband acoustical holography," in *Proceedings of Inter-Noise* (2014), paper 44.

[7]J. Hald, "Extension of acoustic holography to cover higher frequencies," in *Proceedings of Automotive Acoustics Conference (Autoneum)* (2015), paper 20.

[8]J. Hald, "Wide-Band Acoustic Holography," International PCT patent, International publication number WO2015/010850 A2 (2015).

[9]T. Suzuki, "Generalized inverse beamforming algorithm resolving coherent/incoherent, distributed and multipole sources," in *Proceedings of AIAA Aeroacoustics Conference* (2008), paper 2008-2954.

[10]J. Gomes and P. C. Hansen, "A study on regularization parameter choice in near-field acoustical holography," in *Proceedings of Acoustics'08 (Euronoise)* (2008), pp. 2875–2880.

[11]M. Grant and S. Boyd, "CVX: MATLAB software for disciplined convex programming, version 2.1," available at http://cvxr.com/cvx (Last viewed March 2016).

[12]E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic, London, 1999).

[13]M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," J. Res. Natl. Bur. Stand. (U. S.) **49**, 409–436 (1952).

[14]V. Frayssé and L. Giraud, "A set of conjugate gradient routines for real and complex arithmetics," CERFACS Technical Report TR/PA/00/47 (2000).