



Wideband acoustical holography

Jorgen HALD

Brüel & Kjaer Sound & Vibration Measurement A/S, Denmark

ABSTRACT

Patch Near-field Acoustical Holography (NAH) methods like SONAH and ESM are limited to relatively low frequencies where the average array inter-element spacing is less than half a 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 must be used, while beamforming requires a larger distance to limit sidelobe issues. Wideband Holography 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 underlying problem solved is that at high frequencies the microphone spacing is too large to meet the spatial sampling criterion, and thus there is no unique reconstruction of the sound field. A reconstruction must therefore have a built-in “preference” for specific forms of the sound field. Doing just a Least Squares solution will result in reconstructed sound fields with sound pressure equal to the measured pressure at the microphones, but very low elsewhere. By building in a preference for compact sources, a smoother form of the reconstructed sound field is enforced.

Keywords: Noise Source Identification, Microphone arrays, Array signal processing.
I-INCE Classification of Subjects Number(s): 72, 75.7, 74.7.

1. INTRODUCTION

Near-field Acoustical Holography (NAH) is based on performing 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 NAH (SONAH) (2), can work with irregular microphone array geometries, but still require an array element spacing of less than half the wavelength. As described by Hald (3), 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 two to three 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 wide-band 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 Compressive 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. In general, these methods 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 (amplitudes) of most functions equal to zero. The sparse set of expansion coefficients is typically identified by solution of an inverse problem subject to a penalty on the 1-norm of the coefficient vector. Some initial work on CS was published by Tibshirani (4), and the methodology has been widely used in signal analysis applications, see for example Chen (5), and for beamforming as published by, for example, Müjdat (6), Yardibi (7), Edelman (8) and Zhong (9).

Recently, Chardon et al. (10) published an application of CS to acoustic holography on vibrating

plates. They used random and pseudo-random array geometries, and they made use of the fact that plate vibration can often be represented to a good accuracy by relatively few functions of plane-wave form. Suzuki (11) published a similar method, but he used a monopole/dipole point source model of the same type as ESM, and he enforced sparsity in the source model by use of a 1-norm penalty on the solution vector. Suzuki called his method Generalized Inverse Beamforming, focusing thus on rather long measurement distances. Chardon et al. used readily available Matlab convex optimization software to solve the inverse problems including 1-norm minimization, while Suzuki developed his own iterative solver. Antoni (12) used a Bayesian formulation for setting up the inverse problem to be solved for the complex source amplitudes in a point source model. His method supports the specification of piston-like behaviour of the point sources over small areas, denoted as “source coherence”, to alleviate the underdetermined nature of the inverse problem when using an array above its frequency of half-wavelength average element spacing.

The present paper describes a new method called Wideband Holography (WBH), which is covered by a pending patent (13). The method is similar to the method published by Suzuki. However, instead of applying a 1-norm penalty to enforce sparsity in the monopole source model, WBH uses a dedicated iterative solver that enforces sparsity in a different way. The paper is organized as follows: chapter 2 outlines the theory, chapter 3 discusses the applied array design, chapter 4 describes application to some simulated measurements, chapter 5 deals with real measurements, and finally a summary is given in chapter 6.

2. THEORY

Input data for patch holography processing is typically obtained by simultaneous acquisition with an array of M microphones, indexed by $m=1,2,\dots,M$, 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} :

$$\mathbf{G} = \mathbf{V}\mathbf{S}\mathbf{V}^H, \quad (1)$$

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

$$\mathbf{p}_\mu = \sqrt{s_\mu} \mathbf{v}_\mu. \quad (2)$$

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 is a single vector \mathbf{p} with measured complex sound pressure values for all microphones.

WBH uses a source model in terms of a set of elementary sources or wave functions 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). With A_{mi} representing the sound pressure at microphone m due to a unit excitation of monopole number i , the requirement that the modelled sound pressure at microphone m must equal the measured pressure p_m can be written as:

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

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

$$\mathbf{p} = \mathbf{A}\mathbf{q}, \quad (4)$$

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 Compressive Sensing 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. This is done by adding a penalty

proportional to the 2-norm of the solution vector when minimizing the residual norm:

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

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

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

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), ref. Gomes and Hansen (14). 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. As will be shown in section 4.1, 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 therefore be much too low. Another effect is ghost sources because available measured data is far from determining a unique solution.

If the true source distribution is sparse, or close to sparse, the above phenomena can be alleviated by replacing the 2-norm in the penalty term by a p -norm:

$$\text{Minimize}_{\mathbf{q}} \|\mathbf{p} - \mathbf{A}\mathbf{q}\|_2^2 + \theta^2 \|\mathbf{q}\|_p^p \quad (7)$$

where $0 < p \leq 1$, see for example Müjdat (6). An equivalent problem with $p = 1$ was solved by Zhong et al. (9) and Chardon et al. (10):

$$\text{Minimize}_{\mathbf{q}} \|\mathbf{q}\|_1 \quad \text{subject to} \quad \|\mathbf{p} - \mathbf{A}\mathbf{q}\|_2 \leq \delta. \quad (8)$$

With $p = 1$, both of the equivalent optimization problems are convex and can be solved quite efficiently by available Matlab libraries (10). Still, the computational demand is significantly higher than for the Tikhonov problem in Eq. (5) because no analytic solution exists, and also there are no tools available for automated determination of the regularization parameters θ and δ for given input data \mathbf{p} . Use of the p -norm with $0 < p \leq 1$ as in (7) and (8) will have the effect of favouring sparse solution vectors, i.e., source amplitude vectors \mathbf{q} with a maximum number of elements equal to zero.

A main idea behind the WBH method of the present paper is to remove/suppress the ghost sources associated with the real sources in an iterative solution process, starting with the strongest real sources. We define the residual vector \mathbf{r} related to Eq. (4) as:

$$\mathbf{r}(\mathbf{q}) \equiv \mathbf{p} - \mathbf{A}\mathbf{q}, \quad (9)$$

and the quadratic residual function F to be minimized as:

$$F(\mathbf{q}) \equiv \frac{1}{2} \|\mathbf{r}(\mathbf{q})\|_2^2. \quad (10)$$

Denoting by \mathbf{q}_k the approximate solution after the k th iteration step, we compute first the step $\Delta\mathbf{q}_k$ that minimizes the residual function F in the steepest descent direction:

$$\Delta\mathbf{q}_k = s_k \mathbf{w}_k. \quad (11)$$

Here, \mathbf{w}_k is the negative gradient vector:

$$\mathbf{w}_k \equiv \mathbf{A}^H \mathbf{r}(\mathbf{q}_k) = \mathbf{A}^H (\mathbf{p} - \mathbf{A}\mathbf{q}_k), \quad (12)$$

and s_k is the step length to the minimum along that direction:

$$s_k \equiv \frac{\mathbf{g}_k^H \mathbf{r}(\mathbf{q}_k)}{\mathbf{g}_k^H \mathbf{g}_k}, \quad (13)$$

the vector \mathbf{g}_k being defined as:

$$\mathbf{g}_k \equiv \mathbf{A}\mathbf{w}_k. \quad (14)$$

The next candidate solution vector is now:

$$\tilde{\mathbf{q}}_{k+1} \equiv \mathbf{q}_k + \alpha \Delta \mathbf{q}_k, \quad (15)$$

where α is a possible relaxation factor, typically between 0.5 and 1.0. However, just using the step of Eq. (15) will spread the excitation over all point sources of the source model and, as mentioned above, lead to the introduction of ghost sources. These are associated with the real sources, and when using irregular array geometries they will in general be weaker than the strongest real source(s). We can therefore suppress the ghost sources by setting all components in $\tilde{\mathbf{q}}_{k+1}$ below a certain threshold to zero. The threshold T_k is computed as being a number $D_k > 0$ of decibel below the amplitude $|\tilde{q}_{k+1,\max}|$ of the largest element in $\tilde{\mathbf{q}}_{k+1}$:

$$T_k \equiv 10^{\frac{-D_k}{20}} |\tilde{q}_{k+1,\max}|, \quad (16)$$

so the elements $q_{k+1,i}$ of the next solution estimate \mathbf{q}_{k+1} are computed in the following way:

$$q_{k+1,i} = \begin{cases} \tilde{q}_{k+1,i} & \text{if } |\tilde{q}_{k+1,i}| \geq T_k \\ 0 & \text{otherwise.} \end{cases} \quad (17)$$

The dynamic range of retained source amplitudes, D_k , is updated during the iteration in such a way that an increasing dynamic range of sources will be included, typically:

$$D_{k+1} = D_k + \Delta D. \quad (18)$$

In the limiting case when $D_k \rightarrow \infty$ for $k \rightarrow \infty$, the dynamic range limitation is gradually removed. However, the iteration is stopped when:

$$D_{k+1} > D_{\max} \quad \text{or} \quad \|\mathbf{w}_{k+1}\|_2 < \varepsilon \|\mathbf{w}_0\|_2, \quad (19)$$

where D_{\max} is an upper limit on D_k and ε is a small number. Since, in general, the source model may not be able to completely represent the measured pressure data, $F = 0$ may not be achievable. But a minimum of F , characterized by $\|\mathbf{w}\|_2 = 0$, always exists, explaining the use of $\|\mathbf{w}\|_2$ in the stopping criterion. The following values have been found to work in general very well:

$$\mathbf{q}_0 = \mathbf{0}; \quad \alpha = 1.0; \quad D_0 = 0.1 \text{ dB}; \quad \Delta D = 1.0 \text{ dB}; \quad D_{\max} = 60 \text{ dB}; \quad \varepsilon = 10^{D_{\max}/20}. \quad (20)$$

The upper limiting dynamic range D_{\max} can be changed to match the quality of data, but the choice does not seem to be critical. As will be seen in section 4.2.2, $D_{\max} = 60$ dB supports 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. The value of ε could be chosen independent of D_{\max} , but the value in equation (20) has been found to work well.

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 to be removed, ref. Eq. (17). 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. Another aspect is the fact that a minimum number of the point sources of the model will be assigned an amplitude different from zero, enforcing effectively a sparse solution.

After the termination of the above algorithm based on steepest descent directions, a good estimate of the basic source distribution has been achieved. Proceeding with steepest descent directions will typically provide very slow progress, because of ‘‘zigzagging’’. Experience has shown that good progress can be achieved at this point through a single simple scaling step:

$$\mathbf{q}_{k+1} \equiv \beta \mathbf{q}_k, \quad (21)$$

where the real constant β is chosen to minimize the quadratic function $F(\beta \mathbf{q}_k)$. The counter $k + 1$ on this scaling step is assigned to the variable K .

After that, a few Conjugate Gradient iterations without dynamic range limitation can be performed to ensure convergence to a point very close to a minimum of the function F . Usually, the effect on the source model and the modelled sound field is very small, because the primary Steepest Descent algorithm has already reduced F to be close to a minimum, but it ensures that full convergence has been achieved. The stopping criteria used with the conjugate gradient method are similar to those in

Eq. (19), typically:

$$k - K > 10 \quad \text{or} \quad \|\mathbf{w}_{k+1}\|_2 < \varepsilon \|\mathbf{w}_0\|_2 \quad \text{or} \quad \|\mathbf{w}_{k+1}\|_2 \geq \|\mathbf{w}_k\|_2. \quad (22)$$

Here, one of the gradient criteria will often be fulfilled, while the steepest descent algorithm is terminated mostly by the limit on the iteration count. Further details of the WBH algorithm can be found in reference (13).

The WBH algorithm, which enforces a maximum degree of sparsity in the source distribution, has been found to work well at high frequencies, when a suitable array is used at a not too small measurement distance. However, at low frequencies WBH easily leads to misleading results, when two compact source are so close that available data does not support a resolution of the two with beamforming. In that case, the WBH algorithm (and perhaps other sparsity enforcing algorithms) will often identify a single compact source at a position between the two real sources, so the user might be drawing wrong conclusions about the root cause of the noise. Use of the traditional Tikhonov solution of Eq. (6) will in that case typically show a single large oblong source area covering both of the two real sources. An example is given in section 4.2.1. To minimize the risk of misleading results, it is recommended to use the standard Tikhonov solution up to a transition frequency at approximately 0.7 times the frequency of half wavelength average array inter-element spacing, and above that transition frequency switch to the use of WBH.

3. ARRAY DESIGN

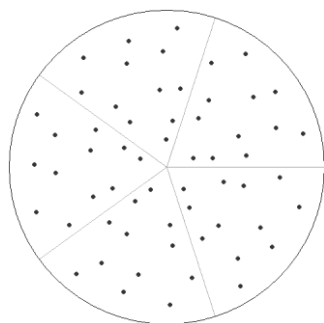


Figure 1 – Geometry of the applied 60-element planar microphone array with 1 m diameter.

As described in the introduction, the method of the present paper follows the principles of Compressive Sensing, being based on measurements with a random or pseudo-random array geometry in combination with an assumed sparsity of the coefficient vector of the source model. The circular, pseudo-random array geometry used in the real and simulated measurements of the present paper is shown in Figure 1. It has 60 microphones uniformly distributed within a diameter of 1 metre, so the average element spacing is approximately 12 cm, implying a transition frequency close to 1 kHz. The geometry, which consists of five identical angular sectors, has been optimized for beamforming measurements up to 6 kHz as described in reference (3): The maximum sidelobe level was minimized subject to the constraint of having the microphones uniformly distributed over the array area. Effectively this

means that the array has an ability to distinguish plane waves from different angles with a minimum of angular leakage, but of course with the angular resolution limit set by the array diameter. This kind of optimization can be performed on the so-called Array Pattern, which represents the directional sensitivity pattern at all frequencies, avoiding the need to optimize the array geometry for specific frequencies. In an application for Direction of Arrival (DOA) estimation, this translates directly into the columns of the Sensing Matrix being optimized for minimum linear dependency (minimum “correlation”, see ref. (15)) between any two of them, which is important for the sparse DOA estimation to work properly. For the present near-field WBH application, the sensing matrix \mathbf{A} is, of course, different from the one used in the far-field DOA application, so it would be interesting to investigate its level of linear dependency. An important finding from simulated measurements with the chosen array design is that the measurement 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. Using the longer measurement distance with WBH has the effect of reducing the difference between the DOA application and the WBH application. Another view of this is the fact that each source in the WBH source model will expose the microphones over a wider area when the measurement distance is increased. To get acceptable low-frequency resolution, however, the distance should not be too long either, so overall the best distance seems to be around twice the average array inter-element spacing.

4. SIMULATED MEASUREMENTS

4.1 Single monopole point source

The aim of this simulated measurement is to demonstrate with a very simple source configuration: 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 using WBH.

A monopole point source is located on the array axis at 28 cm from the array plane, while the source-model mesh is at 27 cm, 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 centred on the array axis, and the source-model mesh is similar, i.e. with 2 cm spacing, but it is extended by 6 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. WBH calculation was performed using dynamic range D_{\max} equal to 40 dB.

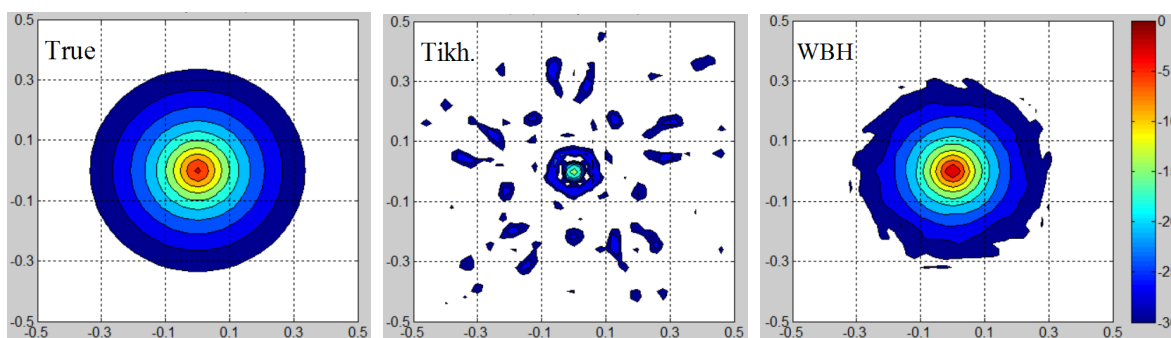


Figure 2 – Contour plots of sound intensity in the “source plane”, 24 cm in front of the array plane. Display range is 30 dB with 3 dB contour interval, and the same scale is used in all three plots.

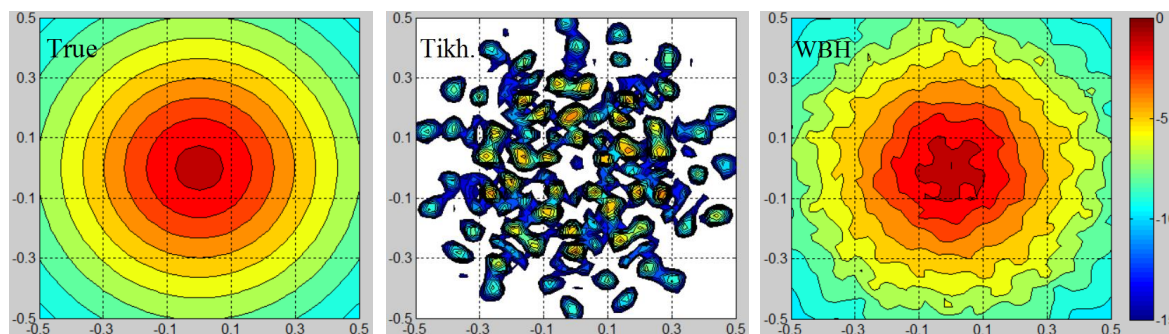


Figure 3 – Contour plots of sound pressure in the array plane. Display range is 15 dB with 1 dB contour interval, and the same scale is used in all three plots.

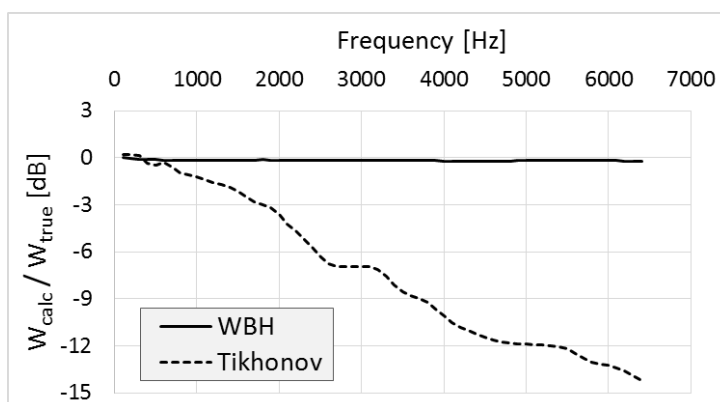


Figure 4 – Relative sound power spectra from the reconstructions based on WBH and on Tikhonov regularization, respectively. Power values are calculated by area integration of plots like those in Figure 2. The sound power from the True intensity map is taken as the reference.

Figure 2 shows the true sound intensity map on the “source plane” at 4 kHz and the corresponding maps calculated from the source model, when Tikhonov regularization is used (Equation (6)) and, when the WBH algorithm is used. The WBH map is very close to the true intensity map, as could be expected in the present case, although the source-model plane is 1 cm from the real monopole point source. The WBH hotspot is a bit more compact with a slightly higher peak than the corresponding one for the true intensity. At the end of the Steepest Descent algorithm in WBH, there were five point sources with non-zero amplitudes in the source model: the one closest to the real source and the four nearest to that. The reconstructed field should therefore have four angular periods around the centre source. Thus, the Conjugate Gradient algorithm must be responsible for introducing the apparent five angular periods in the WBH intensity reconstruction, probably related to the corresponding angular periodicity of the array geometry. 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 the focusing of the radiation on the microphones that can be seen in Figure 3.

Figure 3 shows the true sound pressure level (SPL) on the array plane at 4 kHz and the corresponding sound pressure level generated by the source model, when Tikhonov regularization and WBH, respectively, are used. Looking at the Tikhonov result, it is clear that the 2-norm minimization has used the heavily underdetermined nature of the problem to focus sound radiation towards the microphones to produce a sound pressure close to the measured pressure, while in all other directions the radiated sound is minimized. As can be seen in Figure 4 this means an underestimation of sound power. When WBH is used to obtain source model amplitudes, the reconstructed array-plane SPL is close the true SPL map, although it has some small ripples.

Nevertheless, as shown in Figure 4, the sound power is predicted accurately across the full frequency range, when WBH is used. When Tikhonov regularization is used, sound power underestimation increases quickly with increasing frequency, since the ability of the source model to focus radiation only towards the microphones increases.

4.2 Two coherent in-phase monopole point sources

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 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 centred on the array axis, and the source-model mesh is similar, i.e., with 1 cm spacing, but it is extended by 6 rows/columns in all four directions.

In all the simulated measurements of this section, random noise was added to the complex microphone pressure data at a level 30 dB below the average sound pressure across the microphones.

4.2.1 Equal source levels

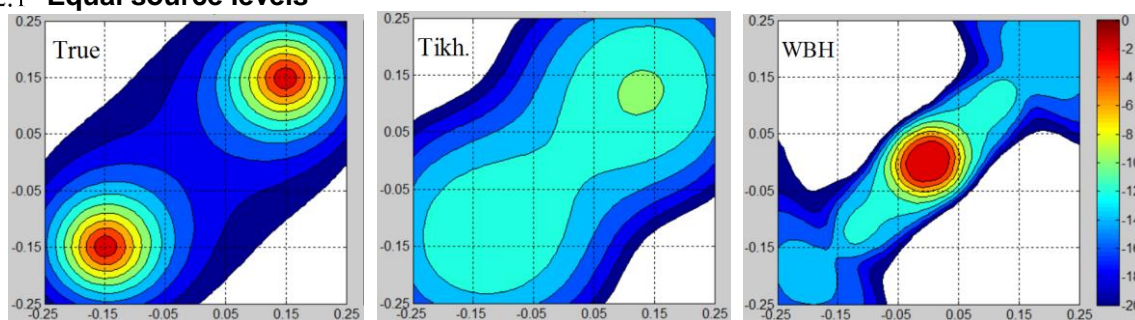


Figure 5 – Contour plots of 400 Hz sound intensity in the reconstruction plane 5 cm from the two point sources. Display range is 20 dB with 2 dB contour interval. The same scale is used in all plots.

First, we consider the case of equal amplitudes of the two monopoles in order to show the advantage of using the Tikhonov regularized solution at the low frequencies. For the WBH calculations, the dynamic range D_{\max} was set to 60 dB. Figure 5 shows sound intensity maps at 400 Hz. From left to right the true sound intensity, the Tikhonov based reconstruction and the WBH reconstruction are shown. Although the Tikhonov result has poor resolution, it indicates quite well that there are two sources. WBH, on the other hand, starts by putting a source at the centre position between the two real sources, and then later during the iteration it will have to add remote model

sources in order for the source model to represent the measured sound pressure data accurately. But the central source stays, which can be misleading.

It is quite easy to show that the first steepest descent direction in Equation (12) is equivalent to a Delay and Sum (DAS) beamforming solution, which has poor low-frequency resolution. The first source will therefore be located at the peak of a DAS map, which in this case is midway between the two real sources. At higher frequencies DAS has good resolution, so therefore the problem of separate sources being replaced by a single central source is probably acceptable above the transition frequency described at the end of section 2.

4.2.2 10 decibel level difference

A main purpose of this section is to demonstrate the ability of the WBH method to identify weak sources in the presence of strong ones. We use the same setup as in section 4.2.1, except that the lower left source is now assigned an excitation 10 dB below that of the upper right source. Figure 6 shows the true and the reconstructed sound intensities at 5 kHz with a 20 dB display range. Again, for the WBH calculations the dynamic range D_{max} was set to 60 dB. Clearly, the two sources are well identified, and the picture looks much the same at all frequencies between 1 kHz and 5 kHz.

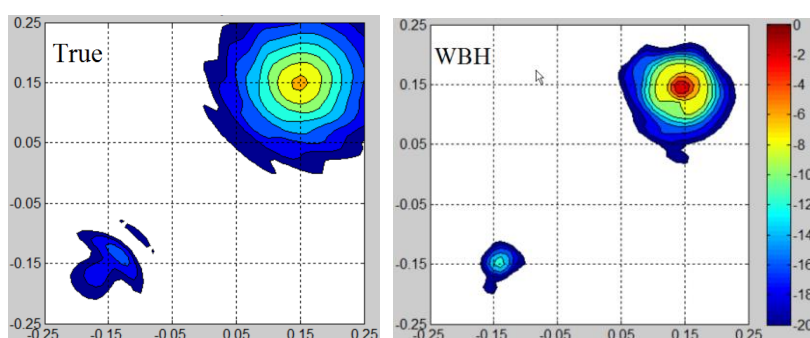


Figure 6 – Contour plots of 5 kHz sound intensity in the reconstruction plane 5 cm from the two point sources. Display range is 20 dB with 2 dB contour interval. The same scale is used in both plots.

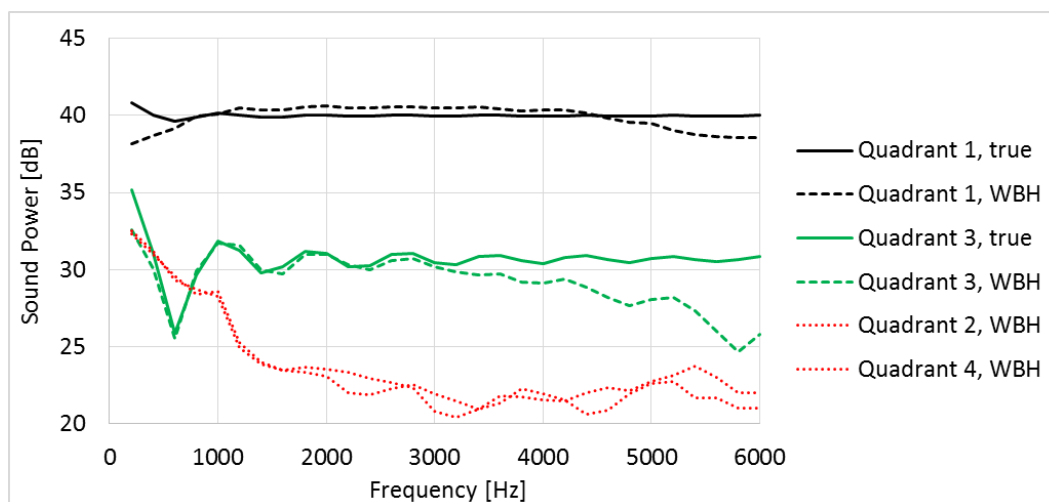


Figure 7 – Sound power spectra obtained by area integration of the true and reconstructed sound intensities over the four quadrants in Figure 6. $D_{max} = 60$ dB in WBH calculation.

Figure 7 shows sound power spectra obtained by area integration of sound intensity maps like those in Figure 6 over the four quadrants, quadrant 1 ($x > 0$ and $y > 0$) containing the strong source and quadrant 3 the weak source. For quadrants 1 and 3 both the true and the WBH reconstruction have been integrated to get reference sound power data for the two sources, while for the remaining two quadrants (2 and 4) only the integrated WBH result is shown. Up to the transition frequency at 1 kHz the Tikhonov regularized solution has been used, while above that frequency, WBH has been applied. The sound power of the strongest source is quite accurately identified over the full frequency range, with a minor underestimation below 500 Hz and above 5 kHz. The sound power of the weak source is very accurate up to 3 kHz. Above that frequency an increasing underestimation is observed, which

remains, however, within 3 dB up to 5 kHz. Above 5 kHz the underestimation seems to increase faster.

The WBH results of Figure 6 and 7 were obtained using the high dynamic range $D_{\max} = 60$ dB. Figure 8 contains sound power spectra like those in Figure 7, the only difference being that a reduced dynamic range, $D_{\max} = 40$ dB, was used for WBH processing. A similar reduction was applied also to the dynamic range used in the Tikhonov regularization at low frequencies. The main effect of the reduced dynamic range is a stronger underestimation of the sound power of the weak source. Thus, to obtain accurate estimation of relatively weak sources, a high value of D_{\max} should be used, typically 60 dB.

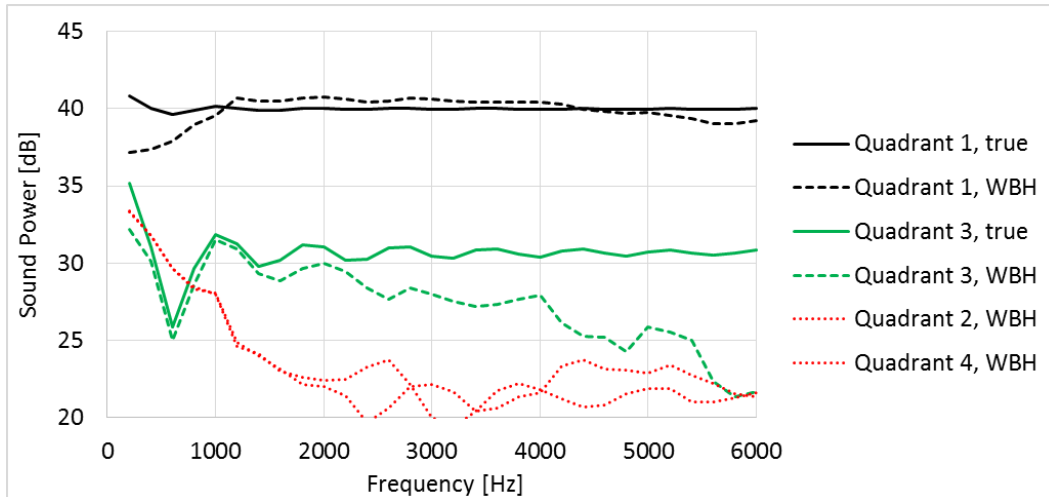


Figure 8 – Sound power spectra obtained by area integration of the true and reconstructed sound intensities over the four quadrants. $D_{\max} = 40$ dB in WBH calculation.

4.3 Plate in a baffle

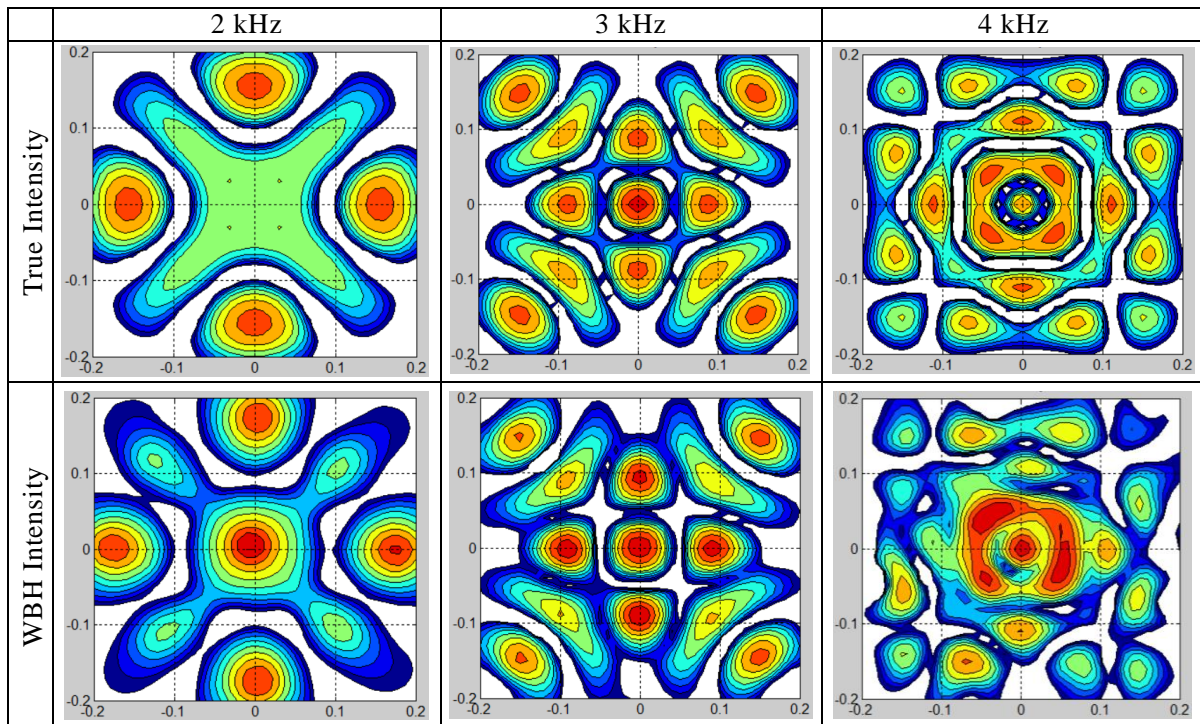


Figure 9 – Contour plots at 2, 3 and 4 kHz of sound intensity in the reconstruction plane 1 cm above the plate. Display range is 20 dB with 2 dB contour interval as in Figure 6. For each frequency, the true sound intensity and the WBH reconstruction use the same scale.

The aim of the simulated plate measurements is to show that the WBH method can give quite good results, even when the true source distribution is not sparse. As an example of a more distributed source, a baffled, centre-driven, simply supported, 6 mm thick, 40 cm \times 40 cm aluminium plate has been used. The coincidence frequency for the plate is at 2026 Hz. The vibration pattern was calculated

using the formulation by Willams (16), 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, and the dynamic range D_{\max} was set to 60 dB.

Figure 9 shows the true sound intensity and the corresponding WBH reconstruction at 2, 3 and 4 kHz with a 20 dB display range. Overall the reconstruction is good, with a bit too high weight on the central area, perhaps because the central sources are first retained during the iterative solution. At 4 kHz the WBH intensity pattern starts getting distorted, because there are many details which are very small compared to the average microphone spacing. The reconstruction accuracy at the highest frequencies can be improved by an increase of the measurement distance, probably because each detail will expose a wider array area. At these frequencies a measurement distance around three times the average microphone spacing has been found to be a good choice, but of course at the expense of slightly poorer low-frequency resolution.

Figure 10 shows the relative sound power spectrum of the WBH reconstruction: At each frequency, the reconstructed and true sound intensity maps (as shown in Figure 9) have been area-integrated, and the ratio between the two sound power values are shown in Figure 10 in decibel. There is a consistent small underestimation, but up to 5 kHz it remains within 2 dB. Above 5 kHz the underestimation increases rapidly.

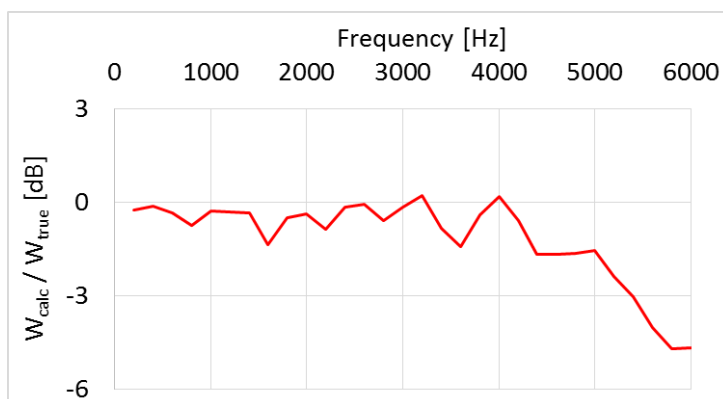


Figure 10 – Reconstructed sound power relative to true sound power in decibel.

5. REAL MEASUREMENTS

5.1 Two mouth simulators

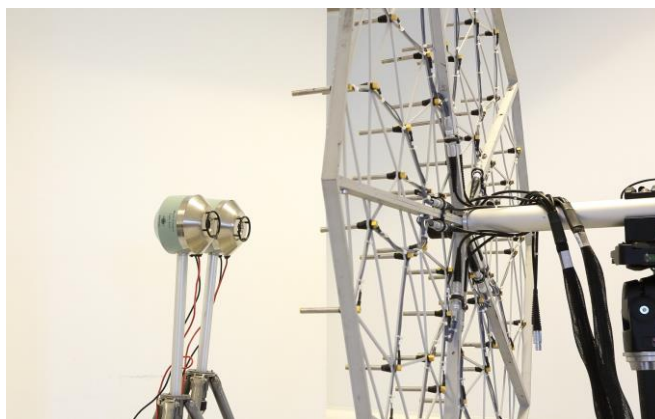


Figure 11 – Setup for measurement on two Brüel & Kjær Mouth Simulators Type 4227.

Figure 11 shows two Brüel & Kjær Mouth Simulators Type 4227 set up 36 cm from the array and with 12 cm separation between the two units. So here, 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 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 $36 \text{ cm} \times 15 \text{ cm}$. The measurements were performed in a normal room with no acoustical treatment.

The array measurement consisted in simultaneously recording 10 seconds of time data with 12.8 kHz bandwidth from all array microphones. As described in the first paragraph of section 2, the processing started with averaging of the 60×60 element cross-spectral matrix between all array microphones. Then, a Principal Component decomposition was performed of that matrix, 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 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 6 rows/columns in all four directions.

Figure 12 shows contour plots of the reconstructed sound intensity for the five 1/3-octave bands at 2, 2.5, 3.15, 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.

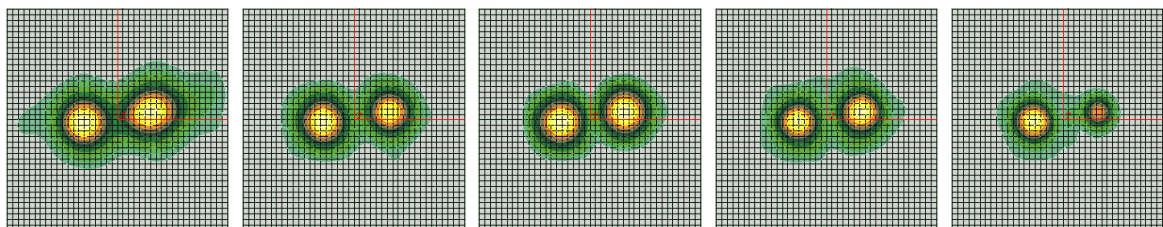


Figure 12 – Contour plots of reconstructed sound intensity in the 1/3-octave bands at 2, 2.5, 3.15, 4 and 5 kHz. Display Range is 20 dB.

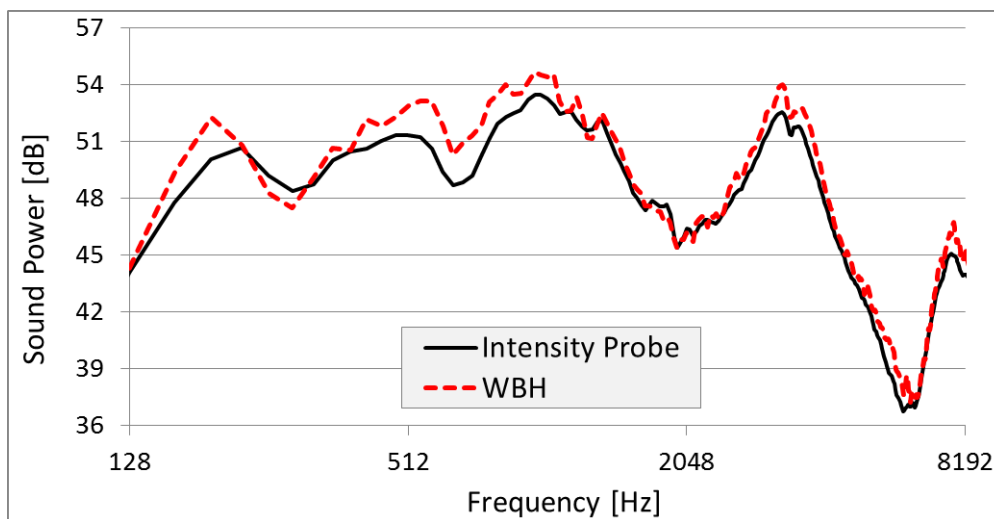


Figure 13 – Comparison of narrow-band sound power spectra from the intensity probe scan and from WBH processing of the array data.

Figure 13 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 Figure 13 should therefore be expected. At the lowest frequencies, there are significant effects of a walls,

floor and ceiling that influence the Tikhonov regularized reconstruction. Apart from that, the overall agreement is very good.

In many practical applications, for example on engines and gearboxes, the source will be far from planar. The sensitivity of the method to sources not being in the assumed/specified source plane is therefore important. This sensitivity is investigated here based on the same array measurement 36 cm from the two mouth simulators, but assuming different source distances in the WBH processing. Assuming, as an example, the source plane to be 46 cm from the array, the reconstruction will be performed in that assumed source plane using a source model 1.5 cm behind the assumed source plane. Figure 14 shows the area-integrated sound power spectra obtained with the assumed source plane being at distances of 26, 31, 36, 41 and 46 cm from the array. The WBH processing (used above 1 kHz) seems, in general, to be less sensitive to variations in the assumed source plane distance than the Tikhonov solution. The only exception is that use of a much too short assumed source distance causes WBH to significantly underestimate the sound power at the highest frequencies. So real sources far behind the assumed WBH source plane are significantly underestimated above, in this case, 5 kHz. Looking at the high sensitivity of the Tikhonov solution at the lowest frequencies, the level variation is close to 5 dB, which is actually equal to $20 \cdot \log(46/26)$. This indicates that roughly a single point source at the assumed distance is given an amplitude to fit the measured level in front of the source. This leads to the amplitude of the point source being proportional to the assumed distance.

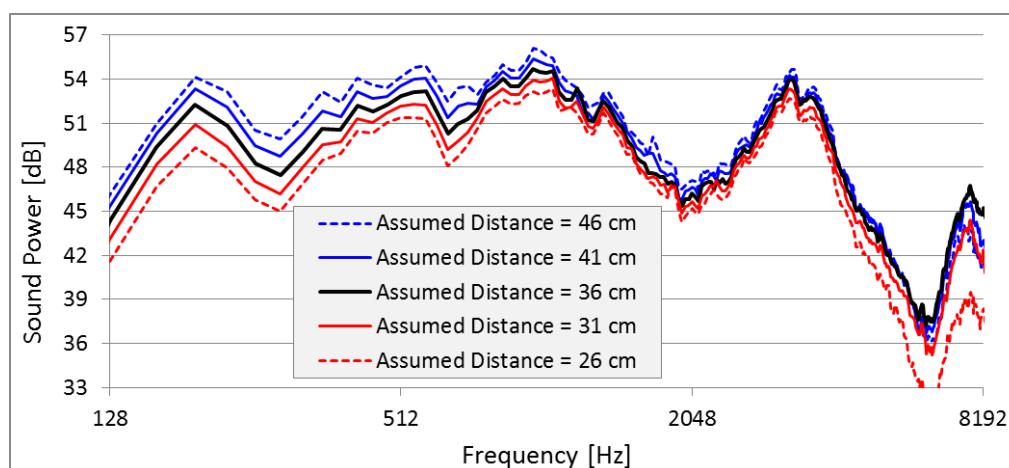


Figure 14 – Comparison of the WBH sound power spectra obtained from the same measurement at a distance of 36 cm by assuming different values of the source distance.

6. CONCLUSIONS

An iterative algorithm has been described for sparsity enforcing near-field acoustical holography over a wide frequency range based on the use of an optimized pseudo-random array geometry. The method, here called Wideband Holography (WBH), can be seen as an example of Compressed Sensing. It was argued, and demonstrated by a simulated measurement, that it is advantageous to supplement the WBH algorithm with a Tikhonov regularized solution at the lowest frequencies. The algorithm has been tested by a series of simulated measurements on point sources and on a plate in a baffle, and subsequently by a real measurement on two small loudspeakers. To check the sound power estimation, a scanned measurement with a sound intensity probe was performed on the same loudspeaker setup. 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 dynamic range. 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. Engine measurements are also characterized by having sources at different distances, and therefore the sensitivity of the WBH algorithm to sources being outside the assumed source plane was investigated. In general the method was found to work surprisingly well with distributed sources, such as vibrating plates, and with sources outside the assumed source plane.

REFERENCES

1. Sarkissian A. Method of superposition applied to patch near-field acoustical holography. *J Acoust Soc Am.* 2005; 118(2):671–678.
2. Hald J. Basic theory and properties of statistically optimized near-field acoustical holography. *J Acoust Soc Am.* 2009; 125(4):2105-2120.
3. Hald J. Array designs optimized for both low-frequency NAH and high-frequency beamforming. *Proc INTER-NOISE 2004.*
4. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society* 1996; 58(1):267-288.
5. Chen SS. Application of basis pursuit in spectrum estimation. *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing* 1998. p. 1865-1868.
6. Müjdat C, Malioutov DM, Willsky AS. A variational technique for source localization based on a sparse signal reconstruction perspective. *Proc IEEE International Conference on Acoustics, Speech and Signal Processing* 2002.
7. Yardibi T, Li J, Stoica P, Cattafesta LN. Sparsity Constrained Deconvolution Approaches for Acoustic Source Mapping. *Proc. AIAA Aeroacoustics Conference* 2008.
8. Edelmann GF, Gaumond CF. Beamforming using compressive sensing. *J Acoust Soc Am.* 2011; 130(4): EL232-EL237.
9. Zhong S, Wei Q, Huang X. Compressive sensing beamforming based on covariance for acoustic imaging with noisy measurements. *J Acoust Soc Am.* 2013; 134(5):EL445-EL451.
10. Chardon G, Daudet L, Peillot A, Ollivier F, Bertin N, Gribonval R. Near-field acoustic holography using sparse regularization and compressive sampling principles. *J Acoust Soc Am.* 2012; 132(3):1521-1534.
11. Suzuki T. Generalized Inverse Beam-forming Algorithm Resolving Coherent/Incoherent, Distributed and Multipole Sources. *Proc. AIAA Aeroacoustics Conference* 2008.
12. Antoni J. A Bayesian approach to sound source. *J Acoust Soc Am.* 2012; 131(4):2873-2890.
13. International patent application no. PCT/EP2014/063597.
14. Gomes J, Hansen PC. A study on regularization parameter choice in Near-field Acoustical Holography. *Proc. Acoustics'08 (Euronoise 2008).* p. 2875-2880.
15. Foucart S, Rauhut H. A mathematical introduction to compressive sensing. Birkhauser Verlag GmbH, Germany: Springer Science; 2013.
16. Williams EG. *Fourier Acoustics.* Academic Press, 1999.