Improving the resolution of three-dimensional acoustic imaging with planar phased arrays

Xenaki, Angeliki; Jacobsen, Finn; Fernandez Grande, Efren

Published in:
Journal of Sound and Vibration

Link to article, DOI:
10.1016/j.jsv.2011.12.011

Publication date:
2012

Improving the resolution of three-dimensional acoustic imaging with planar phased arrays

Angeliki Xenaki[a,*], Finn Jacobsen[a], Efren Fernandez-Grande[a]

[a]*Acoustic Technology, Department of Electrical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark*

Abstract

This paper examines and compares two methods of improving the quality of three-dimensional beamforming with phased microphone arrays. The intended application is the detection of aerodynamic noise sources on wind turbines. Both methods employ Fourier based deconvolution. The first method involves a transformation of coordinates that tends to make the response to a point source, the point spread function, more shift invariant. The result is a significant improvement in sound source imaging in the transformed coordinate system. However, the inverse transformation to Cartesian coordinates introduces range dependent resolution limitations because of the irregular distribution of the focal points. The second method combines the transformation of coordinates with an alternative scanning technique. This method can be used in near field three-dimensional acoustic imaging to produce maps free of sidelobes and with constant resolution. The robustness of the proposed methods is validated both with computer simulations and experimentally.

Keywords:
Delay-and-Sum beamforming, Spatial deconvolution, Dynamic focusing, 3D acoustic imaging

1. Introduction

The beamformer maps produced by phased microphone arrays are usually severely affected by spatial resolution limitations and effects of sidelobes [1]. This distortion may be regarded as a convolution of the distribution of uncorrelated sources and the response of the beamformer to a point source, the point spread function, which in general depends on a number of parameters, including the position of the focal point. Deconvolution can, in principle, reduce the effects of sidelobes and the limited resolution to insignificance [2]. Under certain conditions it is possible to improve the spatial resolution and suppress sidelobe effects using deconvolution methods that can be implemented using spectral (Fourier based) methods. Unlike, e.g., the original DAMAS algorithm [2], such methods require a moderate computational effort [3, 4]. However, spectral deconvolution methods are based on the assumption that the beamformer's point spread function is shift invariant, that is, depending only on the distance between the focal point of the beamformer and the position of the point source and not on the specific positions. Since this is not true in many cases of interest, the deconvolution would yield only approximate results. A coordinate transformation can improve the shift invariance within a specific spatial region and thus improve the results [5, 6]. However, the restricted region of validity imposes limitations to the flexibility of the method [7], and the coordinate transformation required for the implementation of deconvolution by spectral methods implies a range dependent resolution. The purpose of this work is to investigate the principles of the spatial deconvolution method and its applicability in three-dimensional (3D) applications in order to improve the quality of beamformer maps. The paper revises a coordinate transformation to comply with setups in near field aeroacoustic measurements and introduces an alternative scanning method based on dynamic focusing that provides 3D beamforming maps with constant resolution and significantly suppressed sidelobes.

*Corresponding author.
Email address: anxe@elektro.dtu.dk (Angeliki Xenaki)

The remainder of this paper develops as follows: first, in Section 2 the formulation of the conventional delay-and-sum beamforming for near field source imaging is presented. Thereafter, in Section 3, the convolutional formulation of the array response as well as the inverse problem of deconvolution are addressed. In Section 4, the considerations of the coordinate transformation are reported and an alternative focusing technique for accurate 3D sound source imaging is introduced. Next, in Section 5, simulation and experimental results from application of spatial deconvolution techniques on near field delay-and-sum beamforming are compared and discussed. Finally, Section 6 concludes the paper.

2. Beamforming and near field source imaging

Beamforming involves sampling the sound pressure with spatially distributed transducers and post-processing the data by a digital procedure that scans the scene of interest for sound sources. In conventional near field delay-and-sum beamforming, the steering procedure is achieved by applying appropriate individual delays to the pressure signals from the transducers, followed by a summation of all the delayed signals. As a result, signals coming from the point of focus will add up coherently maximising the beamformer output, whereas signals from other positions will be attenuated, leading to a specific beamformer directivity pattern. The individual delays are calculated from the microphone spacing in terms of travel time differences or phase shifts and are adjusted consecutively to focus the array's beam at a particular point in space [1].

Expressed in the frequency domain, the near field delay-and-sum response of a planar array comprising M microphones focused at the point $\mathbf{r}$ is

$$B(\mathbf{r},\omega) = \frac{1}{M}\sum_{m=1}^{M} w_m P_m(\mathbf{r}',\omega) e^{-jk(|\mathbf{r}|-|\mathbf{r}-\mathbf{r}_m|)}, \tag{1}$$

where $P_m(\mathbf{r}',\omega)$ is the sound pressure at the position of the m'th microphone, $\mathbf{r}_m$, due to an assumed point source with strength P_0 at position $\mathbf{r}'$,

$$P_m(\mathbf{r}',\omega) = P_0 \frac{e^{-jk|\mathbf{r}'-\mathbf{r}_m|}}{|\mathbf{r}'-\mathbf{r}_m|}, \tag{2}$$

and w_m is a weighting factor. In the following uniform weighting $w_m = 1$ is considered.

In a stationary sound field the mean square output of the beamformer can be expressed in terms of a sum of cross-spectra, C_{mn}, of the microphone signals [1],

$$\overline{|B(\mathbf{r},\omega)|^2} = \frac{1}{M^2}\sum_{m=1}^{M} C_{mm} + \frac{1}{M^2}\sum_{m\neq n}^{M} C_{mn} e^{-jk(|\mathbf{r}-\mathbf{r}_m|-|\mathbf{r}-\mathbf{r}_n|)}. \tag{3}$$

The beamformer response can be improved if an amplitude correction that accounts for the attenuation of the spherical wave sound field at the various transducers is incorporated. This problem has been addressed by Elias [8] and by Christensen and Hald [9]; their solution is based on minimising an error function between the measured and the modelled cross-spectra with respect to amplitudes. The resulting optimal imaging function is

$$I^2(\mathbf{r},\omega) = \frac{1}{M}\frac{\left|\sum_{m,n=1}^{M} C_{mn} v(\mathbf{r}-\mathbf{r}_m) v^*(\mathbf{r}-\mathbf{r}_n)\right|}{\sqrt{\sum_{m,n=1}^{M} |v(\mathbf{r}-\mathbf{r}_m)|^2 \, |v(\mathbf{r}-\mathbf{r}_n)|^2}}, \tag{4}$$

where v is defined by

$$v(\mathbf{r}) = \frac{e^{-jk|\mathbf{r}|}}{|\mathbf{r}|}. \tag{5}$$

Excluding the autospectra in the diagonal of the matrix in cross-spectral beamforming tends to reduce the sidelobe level [9]. In this case the imaging function takes the form

$$J^2(\mathbf{r},\omega) = \frac{1}{\sqrt{M(M-1)}} \frac{\left|\sum_{m\neq n}^{M} C_{mn} v(\mathbf{r}-\mathbf{r}_m) v^*(\mathbf{r}-\mathbf{r}_n)\right|}{\sqrt{\sum_{m\neq n}^{M} |v(\mathbf{r}-\mathbf{r}_m)|^2 \, |v(\mathbf{r}-\mathbf{r}_n)|^2}}. \tag{6}$$

The results presented in what follows are based on this expression. At this point it should be emphasised that the imaging functions I and J are alternative expressions for the beamformer output B derived to optimise the phased array response for near field beamforming. Finally, it should be mentioned that random, non-redundant array geometries are favourable since they tend to reduce aliasing effects due to periodicities in the array pattern [1, 10].

3. Spatial deconvolution and spectral methods

The 'point spread function', s [3, 4], is the beamformer's squared response to a point source. A collection of uncorrelated point sources at arbitrary positions in the observation grid, $\mathbf{r}'$, gives rise to a mean square beamformer response that can be generally written as

$$\overline{|b(\mathbf{r})|^2} = \sum_{\mathbf{r}'} q(\mathbf{r}') s(\mathbf{r}|\mathbf{r}'), \tag{7}$$

where q is the source strength and $\mathbf{r}$, $\mathbf{r}'$ are the vectors that denote the current focal point and the position of the point source in the grid, respectively. The notation $s(\mathbf{r}|\mathbf{r}')$ describes the point spread function as a function of the grid position of focus $\mathbf{r}$ due to the presence of a monopole with unit strength at the location $\mathbf{r}'$. In the ideal case the point spread function would be a delta function, and the beamformer output would reveal the location and the relative strength of sources within the region of observation directly. Unfortunately, this is never the case since the finite size of the array and the discrete pattern of microphones impose resolution limitations and give rise to sidelobes.

3.1. Convolutional formulation of the phased array response

The point spread function is shift invariant if it depends only on the distance between the current observation point and the source position and not on the individual positions. This property is reflected in the form:

$$s(\mathbf{r}|\mathbf{r}') = s(\mathbf{r}-\mathbf{r}'). \tag{8}$$

This is generally the case only when the source region is small compared with the distance between the array and the source [3, 4]. The property of shift invariance implies by definition that the point spread function retains its pattern irrespectively of the specific source position on the grid.

Under the assumption of a shift invariant point spread function, Eq. (7) can be expressed in a convolutional form,

$$\overline{|b(\mathbf{r})|^2} = \sum_{\mathbf{r}'} q(\mathbf{r}') s(\mathbf{r}-\mathbf{r}'). \tag{9}$$

The convolutional formulation of the beamformer's output makes it possible to perform the calculations in the wavenumber domain making use of a discrete spatial Fourier transform,

$$\overline{|\mathbf{b}|^2} = \mathbf{q} * \mathbf{s} = F^{-1}\left[F\left[\mathbf{q}\right] F\left[\mathbf{s}\right]\right], \tag{10}$$

where $\mathbf{b}$, $\mathbf{q}$, and $\mathbf{s}$ are the matrices with the beamformer output, the source distribution, and the point spread function, F denotes a 3D Fourier transform, and F^{-1} denotes the corresponding inverse transform.

The advantage of performing the calculations in the wavenumber domain is the computational superiority of the Fast Fourier Transform (FFT). However, it should be emphasised that the validity of Eq. (10) is restricted to the case where the sources are incoherent and the point spread function is shift invariant.

3.2. Review of deconvolution algorithms

Deconvolution techniques are widely used in many fields of imaging for 'cleaning' the obtained beamforming maps. Application of spatial deconvolution for the solution of Eq. (9) aims at removing the influence of the array's response function from the obtained data, allowing the source strength distribution to emerge [2]. An investigation of various iterative deconvolution algorithms in aeroacoustic beamforming has been carried out by Ehrenfried and Koop [4]. Several algorithms were compared in terms of computational load, robustness and limitations, using synthesised data for a 2D case.

The tested algorithms were non-negative least-squares solvers aiming to converge towards the actual planar source distribution through an iterative procedure. The investigation included the classical deconvolution algorithms that involve matrix operations to solve Eq. (7) for $q(\mathbf{r}')$ under the implicit constraint of non-negative source strengths. The algorithms were also modified to incorporate spectral procedures, providing the possibility of using a discrete spatial Fourier transform to perform the operations in the wavenumber domain. This tends to reduce the computational effort significantly compared with the classical methods, resulting in a substantial increase of the calculation speed. However, the Fourier based methods require by definition a shift invariant point spread function, and it is difficult to meet this requirement in typical setups in aeroacoustic measurements. One possibility is simply to ignore the variations of the point spread function across the observation grid, which in this case must be restricted to a very limited region. Another solution, proposed by Ehrenfried and Koop [4], involves nesting the Fourier based algorithm into another iterative procedure that accounts for the variations of the point spread function due to the location of the point source within the observation grid. This method is expected to converge to an accurate solution of the source distribution. However, the nesting of iterative procedures slows down the process significantly, and the computational effort becomes comparable to that of the classical deconvolution methods based on matrix operations.

The so-called DAMAS2 algorithm suggested by Dougherty [3] (which is an extension of the original DAMAS deconvolution algorithm introduced by Brooks and Humphreys [2]) seems to be a promising compromise between accuracy and computational efficiency considering the fact that the task involves scanning a 3D region. The DAMAS2 algorithm, which is an FFT-based deconvolution method, gives good results if the assumption of a shift invariant point spread function is met, and it provides the possibility of regularising the process by incorporating a Gaussian spatial low-pass filter to suppress high wavenumber noise, such as noise leaking into the reconstructed map because of microphone phase mismatch or the presence of strong point sources outside the region of interest.

4. Shift invariance through a coordinate transformation

In order to reduce the effect of varying point spread function, an 'unconventional' set of coordinates for defining the focal points on the observation grid is introduced. This coordinate transformation provides the mathematical frame for formulating the deconvolution problem accurately. Such techniques are common practice in underwater acoustic vision performed with digital beamforming [5, 6], and a similar transformation for near field aeroacoustic measurements has been suggested by Dougherty [3]. However, Dougherty's approach is somewhat different, and the intention in underwater applications is to reduce the computational load for real time imaging exploiting the regularity of a phased array rather than improving the beamforming map iteratively. Specifically, the main difference from Dougherty's beamforming coordinates [3] is in the definition of the coordinate w, which defines the range. In the underwater acoustic applications [5, 6] the expression for the beamforming output in a shift invariant form is derived based on a regular microphone array resulting in a different formulation of the beamformer delays. Accordingly, the coordinate transformation applied here is described briefly in the following.

To perform 3D beamforming two angles, ϕ and θ, are required to define the steering directions, as shown in Fig. 1. The angle ϕ is defined as the angle between the vector $\mathbf{r}$, which denotes the location of the focal point, and its projection to the yz plane and, similarly, the angle θ is the angle between the vector $\mathbf{r}$ and its projection to the xz plane [5]. Note that even though the two angles indicate the azimuth and the elevation of a single point in space, their definition differs from the angles in conventional spherical coordinates.

According to simple geometrical considerations, the vector that determines the position of the current focal point in the observation grid is defined by the set of coordinates

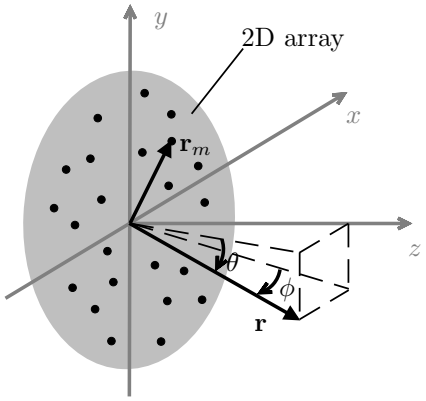

Fig. 1. Steering a beamformer in 3D space.

$$\begin{aligned}(x, y, z) &= r\left(\sin\phi, \sin\theta, \sqrt{\cos^2\phi - \sin^2\theta}\right) \\ &= r\left(u, v, \sqrt{1 - u^2 - v^2}\right),\end{aligned} \tag{11}$$

where the last substitution is based on the transformation

$$u = \sin\phi = \frac{x}{r},\ v = \sin\theta = \frac{y}{r},\ r = |\mathbf{r}|\,. \tag{12}$$

The variables u and v range between -1 and 1, distributing the focal points on a hemispherical surface in the Cartesian coordinate system for constant r.

The transition of the point spread function to a convolutional formulation when it is formulated in terms of the new set of coordinates becomes intuitively clear if one inspects the delays applied to the individual transducers of the array in order to focus the beam at a specific point. Assuming that the microphone array is lying on the xy-plane and the origin of the coordinate system coincides with the geometrical centre of the array, the position of the m'th transducer is defined by the vector $\mathbf{r}_m = (x_m, y_m, 0)$; see Fig. 1. The difference in travel time for a propagating wavefront between a transducer placed at $\mathbf{r}_m$ and another placed at the centre of the array at the position $\mathbf{r}_0 = (0, 0, 0)$ is $(|\mathbf{r}' - \mathbf{r}_m| - |\mathbf{r}'|)/c$ when the beam is focused at a point in the grid at $\mathbf{r}' = (x', y', z')$ that coincides with the correct point source location, and $(|\mathbf{r} - \mathbf{r}_m| - |\mathbf{r}|)/c$ when the beam is focused towards a mismatched grid point at $\mathbf{r} = (x, y, z)$. It follows that the point spread function depends on the quantity

$$c\tau = (|\mathbf{r} - \mathbf{r}_m| - |\mathbf{r}|) - \left(\left|\mathbf{r}' - \mathbf{r}_m\right| - \left|\mathbf{r}'\right|\right). \tag{13}$$

This expression is not a function of $(\mathbf{r}' - \mathbf{r})$. However, replacing the Cartesian coordinates that define the position vectors according to the transformation described by Eq. (11), Eq. (13) expands as

$$c\tau = \left(r\sqrt{\frac{x_m^2 + y_m^2}{r^2} - 2\frac{x_m u + y_m v}{r} + 1} - r\right) - \left(r'\sqrt{\frac{x_m^2 + y_m^2}{r'^2} - 2\frac{x_m u' + y_m v'}{r'} + 1} - r'\right). \tag{14}$$

Applying a second-order binomial expansion of the square root and retaining only the first two terms of the series, Eq. (14) takes the form,

$$c\tau \approx -x_m(u - u') - y_m(v - v') + \frac{1}{2}(x_m^2 + y_m^2)\left(\frac{1}{r} - \frac{1}{r'}\right). \tag{15}$$

Equation (15) is based on the Fresnel approximation, which holds under certain assumptions that in turn define the region of validity. The conditions that determine the region of validity of the presumed approximations are explicitly described in [7]. The last term in Eq. (15) accounts for effects of the curvature of the spherical waves and defines the range resolution that is mainly depending on the dimensions of the array [1]. By introducing the variable $w = r_{\min}/r$ instead of the range term in Eq. (15), the expression assumes the shift invariant form

$$c\tau \approx -x_m(u-u') - y_m(v-v') + \frac{1}{2}\frac{x_m^2 + y_m^2}{r_{\min}}(w-w'), \tag{16}$$

where $r_{\min}$ is the minimum radial distance within the region of validity. This constant is employed in order to normalise the values of the w-variable, which evidently assumes non-negative values up to unity.

The resulting coordinate system is described by the set of dimensionless coordinates

$$(u, v, w) = \left(\sin\phi, \sin\theta, \frac{r_{\min}}{r}\right),\ r = |\mathbf{r}|\,. \tag{17}$$

The steering directions are regularly distributed in the new coordinate system, which implies that they are equally spaced in the sines of the steering angles and the reciprocal of the distance.

The regular grid in the *uvw*-coordinate system is transformed into an irregular grid in the Cartesian coordinate system *xyz*, which is arranged in concentric hemispherical surfaces. The corresponding grid points in the two systems are presented in Fig. 2 for r=1 m and r=2 m. The figure demonstrates that a planar regular grid in the *uvw*-domain corresponds to a hemispherical grid with a higher density of points around the pole in the *xyz*-domain. Planes closer to the origin in the *uvw*-domain are transformed into hemispheres with longer radius in the *xyz*-domain. On the basis of this reciprocal relation, it is often useful to present the w-axis inverted.

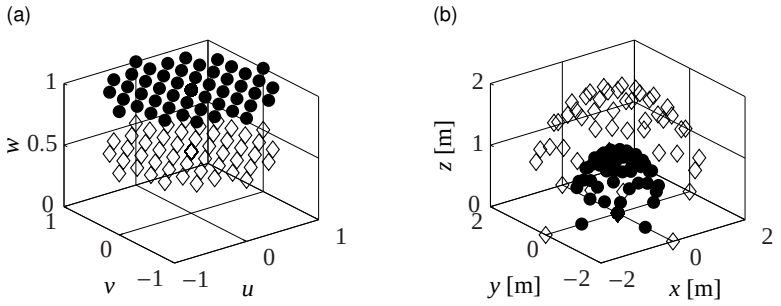

Fig. 2. (a) Regular grid planes in *uvw* coordinate system. (b) The same grid planes inverse transformed into concentric hemispheres in the *xyz* coordinate system. The filled circles represent the case of r=1 m, and the empty diamonds represent the case of r=2 m.

4.1. Dynamic focusing

Solving the spatial deconvolution problem for a 3D region in the wavenumber domain, a 3D FFT algorithm is required to calculate the spatial discrete Fourier transform and its inverse. It is thus required that the observation grid is regularly spaced in the *uvw*-coordinate system and that the Fourier transform is calculated at fixed points. However, a regular grid in the *uvw*-domain results in an irregular arrangement of points in the *xyz*-domain after the inverse transform, such that the focal points are distributed on hemispherical surfaces of gradually increasing radius. Furthermore, the regularity of the grid as defined in the *uvw*-domain implies a gradually sparser distribution of points in the *xyz*-domain as hemispheres with larger radius are considered; see Fig. 2. This has a negative influence on the accuracy of the reconstruction of the beamformer output in the Cartesian coordinate system, necessitating interpolation, which leads to spatially smeared results for distant locations in all dimensions.

A new method is proposed for keeping the density of the points approximately constant in the grid after the inverse transform irrespective of the range. By neglecting the curvature of the hemispherical distribution within the restricted angular segment imposed by the region of validity, each plane parallel to the array is treated separately and finally the results are aggregated to form the 3D imaging map. This method is inspired by a technique of continuous scanning introduced by Murino and Trucco to enhance the depth in underwater acoustic vision [11, 12].

The observation grid, as defined in the Cartesian coordinate system, is separated into *xy*-planes equidistantly spaced, and each plane is processed individually. For each *xy*-plane, a new 2D grid equidistantly spaced in the new coordinate system is defined, and the procedures of delay-and-sum beamforming, spatial deconvolution with 2D FFT and inverse transform into a regular planar distribution defined in the Cartesian coordinate system through interpolation are applied. The coordinate transformation is based on the simplifying assumption that $z \approx r$ all over each parallel plane. This is an acceptable approximation when the examined segment confines the beamforming points close to the pole of the hemispheres.

After processing all the predefined *xy*-planes with a plane-by-plane 2D spatial deconvolution method, the results are aggregated to form an image of the 3D scene of interest. Contrary to the 3D spatial deconvolution methods, this approach allows redefinition of the focusing spatial span individually for each *xy*-plane. Consequently, every *xy*-plane comprises an equal number of focusing points by constraining them within a more limited angular region as more distant *xy*-planes are examined. Since the density of the focal points after the inverse transform remains constant irrespective of the range, the interpolation scheme reproduces a regular grid of points in the *xyz*-domain with the same accuracy.

Figure 3 illustrates the two alternative focusing techniques. The first, referred to as 3D beamforming in the following, assumes a predefined 3D grid in the *uvw*-domain to perform the beamforming processing with spatial deconvolution methods. The second, referred to as dynamic focusing, assumes that the processing is applied consecutively in each predefined parallel plane in 2D. Thus a 2D grid in the *uvw*-domain is redefined for each plane separately. The more distant the examined plane is, the more confined the assigned 2D grid becomes.

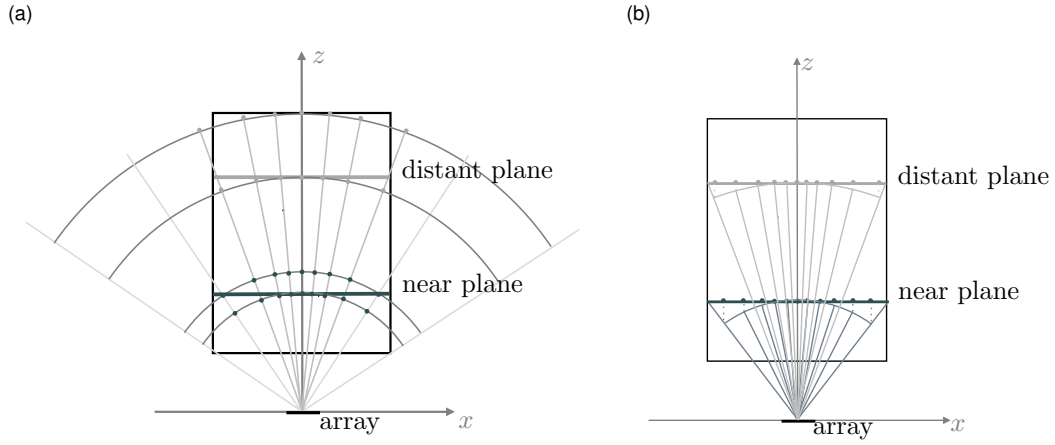

Fig. 3. Distribution of focal points in (a) 3D beamforming and (b) dynamic focusing beamforming.

5. Implementation and application of the method

To examine the performance of the methods described in the foregoing a computer model has been developed. The assumed configuration is a scaled version of an actual setup for wind turbine measurements.

5.1. Model implementation

In noise measurements on wind turbines a planar array is lying on the ground at some distance from the foundation of the wind turbine. The observation plane of main interest is the plane of the airfoil, which is perpendicular to the array. This, in combination with the fact that the airfoil blades may deflect from the perpendicular plane of the rotor during the operation of the wind turbine, indicates the need for effective source imaging on a 3D observation grid. The defined observation grid should circumscribe at least the length of a blade on a perpendicular plane and the extent of the deflection on a parallel plane to the array. On the other hand, the constraints imposed by the attempt to achieve

a convolutional formulation of the point spread function confine the observation grid in a specific spatial region of validity. In practice this means that rather large arrays are required for actual field measurements on wind turbines.

In the case under study a scaled version of such a configuration is used. In addition, the total duration of recording the sound field is supposed to be short enough to assume that the blade remains static. Additionally, it is assumed that the ground is flat and totally absorbing (or completely reflecting). Moreover, the size of the microphones is supposed small enough and their distribution sparse to render the array acoustically transparent. Furthermore, no effects of atmospheric propagation, such as absorption or refraction of energy, are included.

In order to avoid wrap-around effects due to circular convolution inherent in digital signal processing [13] a zero-padded extension of the scene of interest is incorporated to the observation grid. The resulting buffer zone augments the observation grid circumferentially by at least half the length in every dimension. In order to exploit the whole region of validity, which is rather limited, the buffer zone can be extended to a segment broader than 18° if it can be assumed that there are no sources within this zone. In the following simulations, such a zero-padded 3D observation grid comprising 41 points in every dimension, as demonstrated in Fig. 4, is scanned for sound sources. The considered grid has an adequate extent of buffer zone in the Cartesian coordinate system, which also ensures adequate zero-padding of the corresponding grid in the transformed *uvw* set of coordinates allowing comparable results and a uniform presentation.

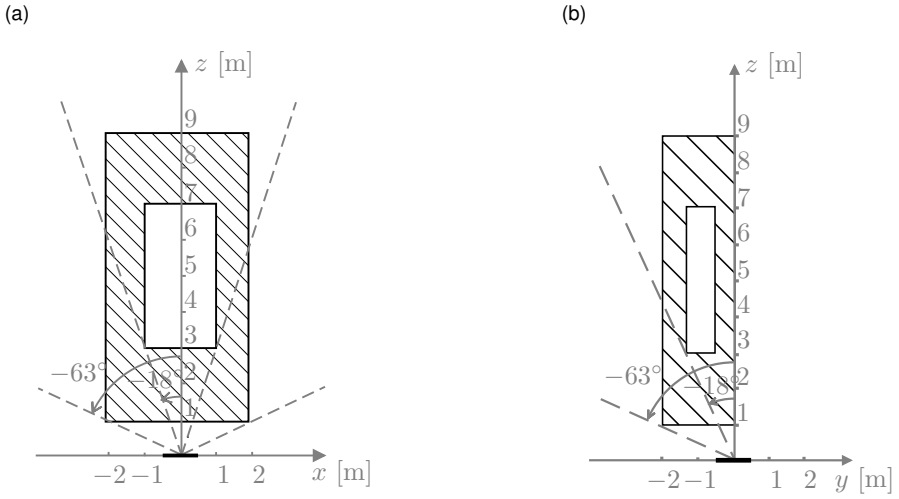

Fig. 4. Observation grid with buffer zone.(a) Intersection at y=0 and (b) intersection at x=0.

The contributions to the aerodynamic wind turbine noise generated at different parts of the rotor airfoil can be assumed to be incoherent. Furthermore, the sources are rather broadband and not highly directional, and they can therefore be modelled as uncorrelated monopoles expected to appear along the leading and the trailing edge of the airfoil [14, 15]. In the implemented model two pairs of uncorrelated point sources are considered within the observation grid in order to examine both the lateral and the range resolution. The same strength is assumed for all sources. Figure 5 shows the distribution of the four sources within the zero-padded observation grid in the conventional Cartesian coordinate system (Fig. 5(a)) and in the transformed one (Fig. 5(b)).

According to the coordinate transformation described by Eq. (17), the location of the sources in the *uvw*-coordinate system is determined by the range in which they appear in the Cartesian coordinate system. Thus even though all sources lie in the same *xz*-plane they will appear in different *uw*-planes. Similarly, the source pairs S1, S2 and S3, S4 that are set to the same *yz*-planes will appear in different *vw*-planes in the transformed coordinate system.

The resulting setup considered in the following simulations is depicted in Fig. 6. The arrangement consists of a planar aperture with a diameter of 1 m with 60 sparsely distributed microphones in a pseudorandom pattern and two pairs of uncorrelated point sources. The size of the microphone array position combined with the frequency range of concern set the limits for near field beamforming in the transformed set of coordinates. In what follows a frequency of 6000 Hz is considered, corresponding to a fairly large near field region with the given array aperture. The origin

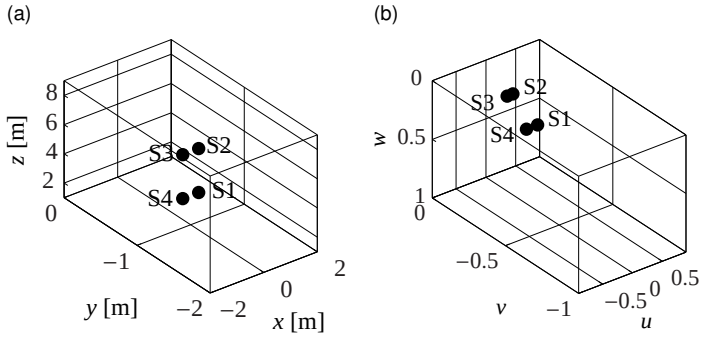

Fig. 5. Source distribution in (a) the *xyz*-coordinate system and (b) the *uvw*-coordinate system.

of the coordinate system is identified at the geometrical centre of the array. The two pairs of sources are positioned at two different range distances, namely 3 and 6 m respectively. The sources of each pair are symmetrically spaced 0.3 m from the origin of the coordinate system on the *x*-axis and shifted 1 m towards the negative side of the *y*-axis to comply with the considerations of an actual measurement setup.

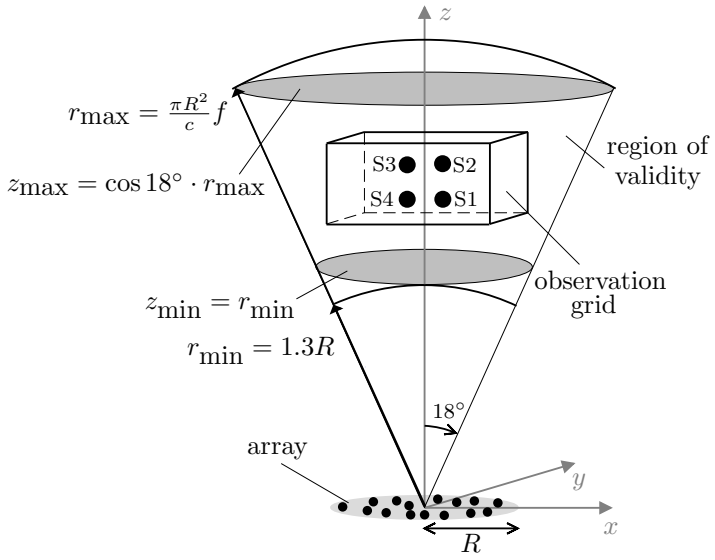

Fig. 6. Test configuration with a planar sparse array. Four omnidirectional sources of equal strength are placed within the observation grid in 3D. The region of validity is the space where the point spread function assumes a convolutional formulation in the transformed set of variables within the array's near field region.

5.2. Pattern of the point spread function

The translational characteristics of the point spread function both in the conventional and the transformed set of coordinates are illustrated in Fig. 7.

A comparison of the beamformer's response to different point source locations in the transformed coordinates reveals that the point spread function retains its form irrespectively of the position of the point source in the observation grid (Figs. 7(c), 7(d), 7(g), 7(h)). Conversely, when beamforming is performed in a grid defined in Cartesian coordinates the point spread function alters its characteristics: the main lobe becomes wider due to the transition of the source to more distant positions and tilted in the perpendicular plane due to transition of the source to off-axis positions (Figs. 7(a), 7(b), 7(e), 7(f)). Consequently, the point spread function has a more shift invariant form in the

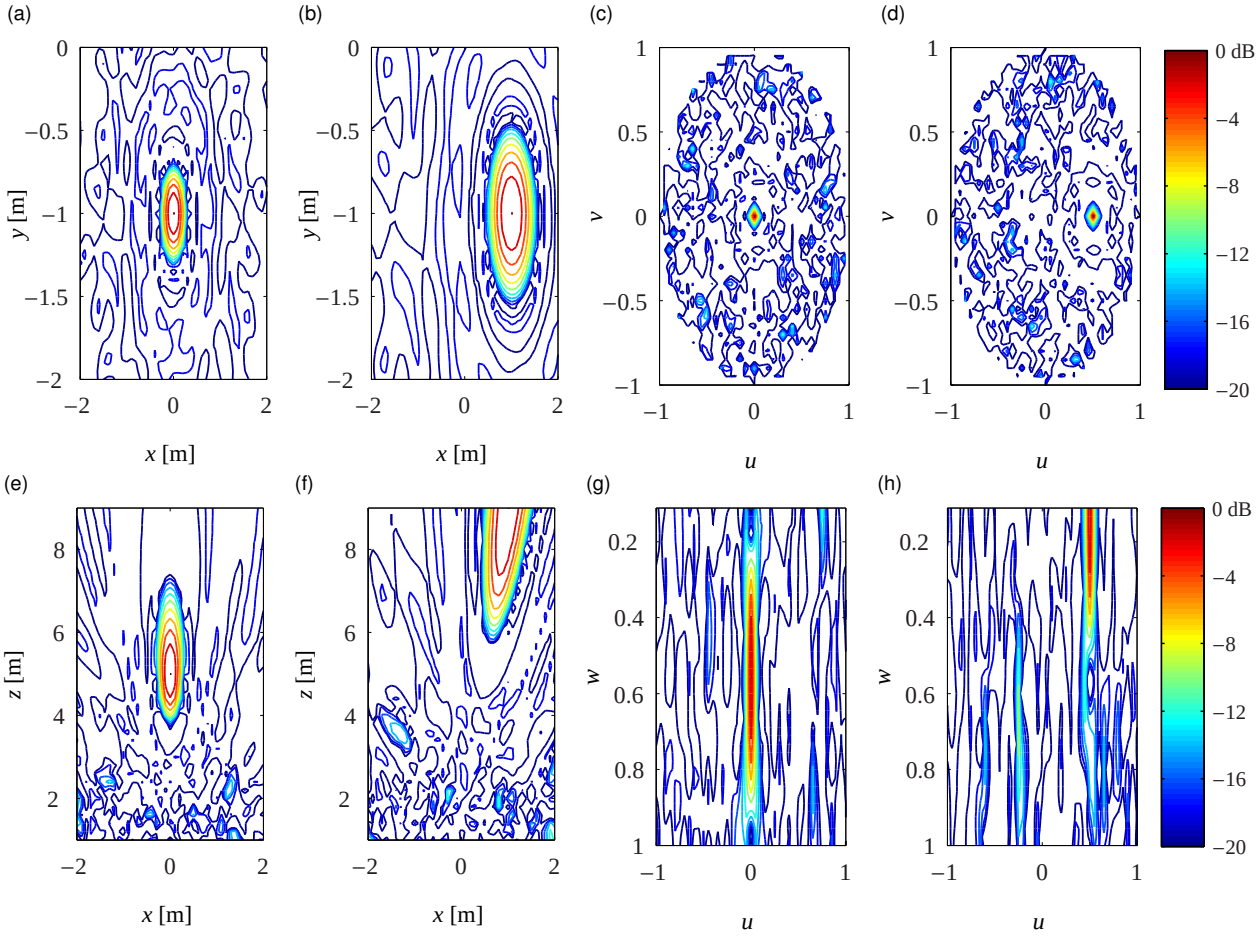

Fig. 7. Comparison of the point spread function at 6 kHz in the two different coordinate systems, *xyz* ((a),(b),(e),(f)) and *uvw* ((c),(d),(g),(h)) for a centred point source position ((a),(c),(e),(g)) and for a random point source position ((b),(d),(f),(h)) in a plane parallel to the array ((a)-(d)) and in a plane perpendicular to the array ((e)-(h)).

new set of coordinates. The details of the translational characteristics of the point spread function in the two coordinate systems are better illustrated by a cross section of the point spread function. In Fig. 8 the point spread function is plotted along x (Fig. 8(a)) and along u (Fig. 8(b)) for two different source positions in the grid.

5.3. Simulation results

The 3D observation grid is scanned for sound sources through different beamforming methods and the simulation results are depicted in Fig. 9. The intersectional plane of interest, perpendicular to the array at $y = -1$ m, is chosen to simplify the comparison. The results are presented in the Cartesian coordinate system, scaled in decibels with reference to the maximum value in the map, and the level is indicated by contour lines with a step of -2 dB on the colour scale. The convergence of the iterative deconvolution procedure indicated that 100 iterations are adequate in this case. Furthermore, no Gaussian regularisation filter is introduced in the deconvolution algorithm since the results are very tolerant to noise due to microphone phase mismatch and/or noise due to the existence of extraneous noise sources.

First, the beamformer output is calculated by applying a conventional delay-and-sum beamforming procedure in an equidistantly spaced observation grid in the Cartesian coordinate system. Figure 9(a) shows that the resulting map is characterised by side lobe contamination and poor resolution. Even though the locations of the sources within the observation grid are indicated, the sources appear spatially smeared and their strengths are not correctly reconstructed. It is apparent that the further from the array the source is, the poorer the resolution becomes. This

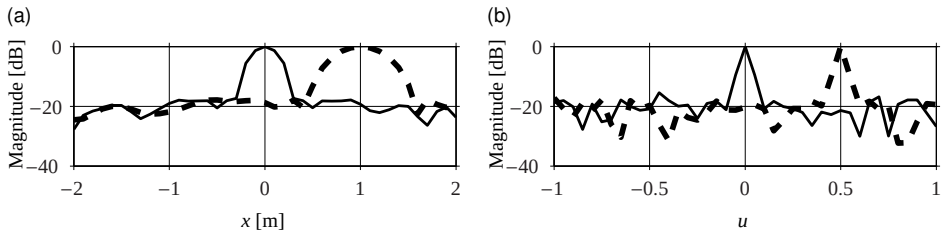

Fig. 8. Translational characteristics of the point spread function (a) as a function of x in the xyz coordinate system and (b) as a function of u in the uvw coordinate system for a centred point source position (solid line) and for a random point source position (dashed line).

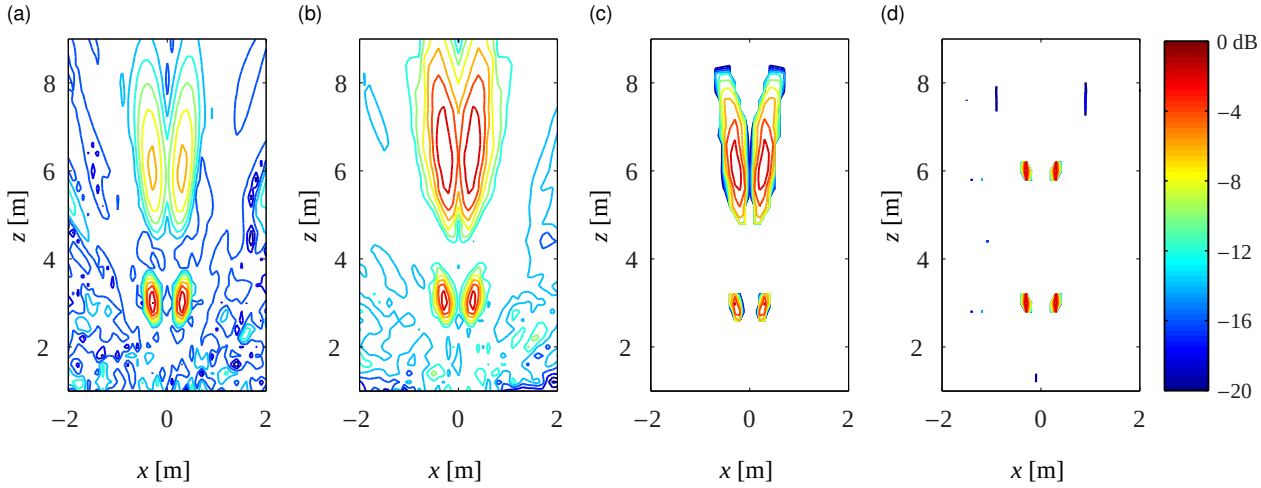

Fig. 9. Simulation results of the beamformer output at 6 kHz after different post-processing techniques: (a) application of the conventional delay-and-sum procedure in Cartesian coordinates; (b) application of the conventional delay-and-sum procedure in the uvw-domain followed by an inverse transform to the xyz-domain; (c) application of the DAMAS2 algorithm in the uvw-domain followed by an inverse transform to the xyz-domain; (d) application of the DAMAS2 algorithm combined with dynamic focusing in the uvw-domain followed by an inverse transform in the xyz-domain.

observation is consistent with the translational characteristics of the point spread function in the Cartesian coordinate system. Next, the same conventional delay-and-sum beamforming procedure is applied in an equidistantly spaced grid in the uvw-domain. In this domain the point spread function assumes a shift invariant form, and therefore the resolution is expected to be constant irrespective of the position of the source in the grid. Eventually, the obtained results should be inversely transformed to the conventional Cartesian coordinate system in order to achieve a more comprehensive illustration of the source distribution. Since the inverse transform of a regular grid in the uvw-domain results in an irregular grid of points in the xyz-domain, interpolation is used to infer the values of the intervening points. Figure 9(b) shows the corresponding beamformer map. The resolution is already improved by performing the delay-and-sum procedure in the uvw-domain. The gradually increasing spatial smearing with increasing range after the inverse transform is attributed to the increasing sparsity of the irregular arrangement of the focal points, which degrades the accuracy of interpolation. Furthermore, the presence of sidelobes is equally prominent with the case of the conventional beamforming procedure in the Cartesian coordinate system. In order to 'clean' the beamformer map, the DAMAS2 deconvolution algorithm is applied. The problem is solved in the uvw-domain in order to fulfill the prerequisite of shift invariant point spread function. The results are inversely transformed to the Cartesian coordinate system, and the corresponding reconstructed source distribution is shown in Fig. 9(c). In the resulting map, the disturbing sidelobe effects have completely vanished. However, even though the beamformer map is dramatically improved, sources at distant positions to the array appear spatially smeared, due to the increased sparsity of the focal points. To compensate for this effect, the dynamic focusing technique is introduced. Figure 9(d) shows the

beamformer output obtained by scanning the grid, as defined in the *uvw*-coordinate system, with dynamic focusing, applying the DAMAS2 deconvolution algorithm and inverse transforming the results in the Cartesian coordinate system. In this map, the source locations and strengths are accurately reconstructed, allowing unambiguous identification of the existing sources. The minor sidelobe contamination is attributed to the presence of other off-plane sources within the 3D grid.

5.4. Experimental results

The predictions presented above indicate that the application of spatial deconvolution in delay-and-sum beamforming in combination with the dynamic focusing technique improves the resolution in 3D sound source imaging significantly. In order to validate these predictions and investigate their robustness in practical applications, experimental measurements were conducted in a large anechoic room at the Technical University of Denmark.

In the experimental setup, four Brüel & Kjær type 4295 omnisource loudspeakers served as sound sources. The loudspeakers were driven with incoherent broadband noise signals of the same level. The array comprised 60 microphones in an optimised geometry that ensures a low sidelobe level up to 6 kHz. A 10 s recording was stored for each of the 60 channels. The cross-spectra were calculated by the *Matlab* function *CPSD*. The time recordings were divided into 160 non-overlapping sequences to ensure acceptable incoherence, and a Hamming window was applied to each sequence. The FFT algorithm over 1024 points and a sampling frequency of 16384 Hz, equal to the one of the data acquisition system, was used to calculate the frequency spectrum of each sequence. Averaging over all sequences and considering all possible microphone combinations resulted in an estimate of the cross-spectral matrix with 513 frequency lines. Eventually, the delay-and-sum procedure and the deconvolution algorithm were applied to map the sound field and reconstruct the distribution of source strengths.

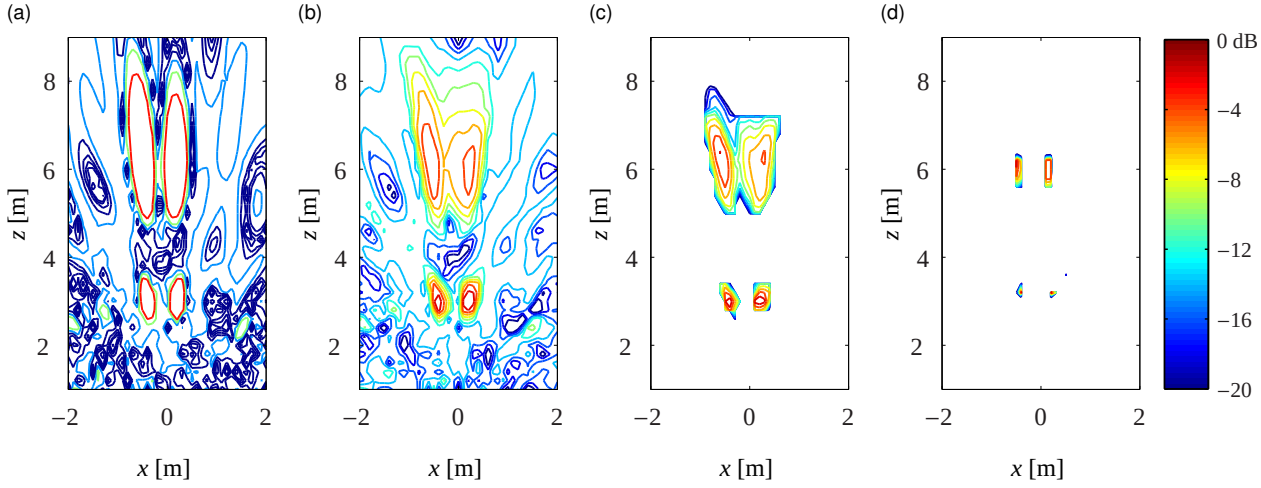

Fig. 10. Experimental results of the beamformer output at 6 kHz after different post-processing techniques: (a) application of the conventional delay-and-sum in Cartesian coordinates; (b) application of the conventional delay-and-sum in the *uvw*-domain followed by an inverse transform to the *xyz*-domain; (c) application of the DAMAS2 algorithm in the *uvw*-domain followed by inverse transform to the *xyz*-domain; (d) application of the DAMAS2 algorithm combined with dynamic focusing in the *uvw*-domain followed by an inverse transform to the *xyz*-domain.

In accordance with the simulation procedure, the beamformer output was first calculated by a conventional delay-and-sum procedure; the resulting map is shown in Fig. 10(a). Next, the coordinate transformation described in the foregoing was applied, and the beamformer output was calculated once more by the delay-and-sum procedure. Figure 10(b) presents the results after the inverse transform and interpolation to form a regular observation grid in the Cartesian coordinate system. In addition, the DAMAS2 deconvolution algorithm was applied in the *uvw*-domain followed by the inverse transform to the Cartesian coordinate system. Figure 10(c) shows the resulting map of the reconstructed source distribution in this case. Finally, the beamformer map was calculated by processing the experimental data with dynamic focusing and the DAMAS2 deconvolution algorithm in the *uvw*-domain. Figure 10(d) shows the reconstructed source distribution after the inverse transform to the Cartesian coordinate system. A comparison of the maps presented in Fig. 10 and the corresponding ones from the simulations presented in Fig. 9 shows

that the experimental results are in agreement with the predictions. The beamformer output calculated by the classical delay-and-sum procedure is characterised by poor resolution and pronounced sidelobe contamination. A distinct improvement of the resolution is achieved by application of a coordinate transformation that attributes shift invariant characteristics to the point spread function and processing the obtained data in the introduced domain. Further processing with a spatial deconvolution method improves the source imaging significantly, resulting in maps totally free of sidelobe effects. Scanning the space of interest with the dynamic focusing technique alleviates the adverse effects of the coordinate transformation on the resolution. The resulting maps depict the sound source distribution with great accuracy and good resolution.

5.5. Improvement of range resolution

The range resolution refers to the sensitivity of the beamformer to distinguish between sources in adjacent *xy*-planes parallel to the array. The improvement of the range resolution after application of the dynamic focusing technique is due to the fact that this method postulates 2D beamforming eliminating the dependence of the point spread function on the range term that is the last term in Eq. (16). Application of the coordinate transformation in 3D beamforming results in a grid defined by hemispheres of gradually increasing radius, and therefore the range distance of the focal points also gradually increases. By applying dynamic focusing, parallel planes remain equidistantly spaced, and hence sources that occur in distant neighbouring planes parallel to the array are not merged. Figure 11 illustrates the limitations in range resolution of the beamformer after the two different processing procedures.

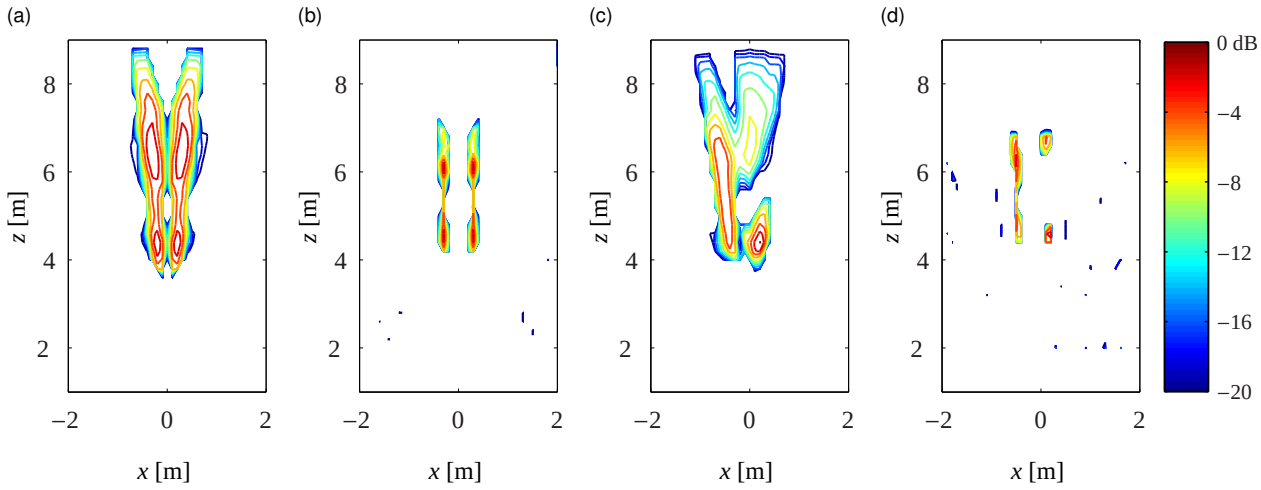

Fig. 11. Improvement of the range resolution in 3D beamforming with spatial deconvolution ((a),(c)) by dynamic focusing ((b),(d)). (a),(b): simulation results; (c),(d): experimental results at 6 kHz. Actual source locations: S1 $(0.3, -1, 4.5)m$, S2 $(0.3, -1, 6)m$, S3 $(-0.3, -1, 6)m$, S4 $(-0.3, -1, 4.5)m$.

Sound sources become indistinguishable by 3D beamforming when they occur in distant *xy*-planes that approach less than 1.5 m (Figs. 11(a), 11(c)). However, by application of dynamic focusing the sources are resolved as separate at the expense of a minor sidelobe contamination due to the presence of off-plane involved in the 2D deconvolution approach (Figs. 11(b), 11(d)).

5.6. Improvement of lateral resolution

The lateral resolution refers to the ability of the beamformer to distinguish between sources in adjacent lateral locations in the same *xy*-plane. Dynamic focusing essentially retains the same density of focal points for each parallel plane after the inverse transform, irrespective of the range of the *xy*-plane considered. Therefore, the lateral resolution remains constant. Figure 12 depicts this improvement both on simulation and experimental results. The lateral distance of the sources is reduced to the half of the initial considerations leaving them only 0.3 m apart.

The lower density of points for distant ranges in 3D beamforming due to the transformation of variables results in a more coarse representation of the beamforming maps in the Cartesian coordinate system. Experimental results

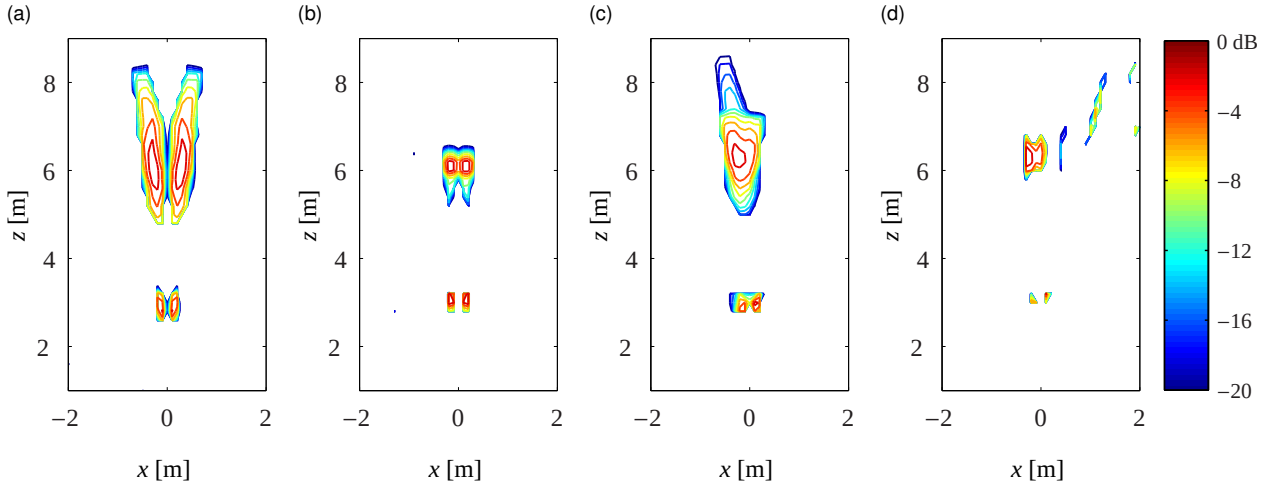

Fig. 12. Improvement of the lateral resolution in 3D beamforming with spatial deconvolution ((a),(c)) by dynamic focusing ((b),(d)). (a),(b): Simulation results; (c),(d): experimental results at 6 kHz. Actual source locations: S1 $(0.15, -1, 3)m$, S2 $(0.15, -1, 6)m$, S3 $(-0.15, -1, 6)m$, S4 $(-0.15, -1, 3)m$.

show that distant sources existing in locations characterised by small lateral separation are not resolved as separate (Fig. 12(c)). Dynamic focusing overcomes this limitation by gradually confining the grid in the *uvw*-domain in narrower angular sections assigning to each parallel plane the same number of points after the inverse transform, and hence keeping the lateral resolution constant irrespective of the range (Fig. 12(d)).

6. Conclusions

The possibility of 'cleaning' a beamformer map by spatial deconvolution in combination with an alternative focusing technique for improving the quality of sound source imaging has been investigated both by simulations and experiments. The classical delay-and-sum beamforming method for scanning a 3D region for sound sources produces rather coarse imaging maps that do not provide the possibility of a straightforward interpretation. The beamforming map resulting after application of the delay-and-sum method in the *uvw*-domain and inverse transformation to the *xyz*-domain is already improved in terms of resolution compared with the one obtained by the classical delay-and-sum beamforming method directly in the *xyz*-domain. Application of deconvolution improves the imaging capabilities of the beamformer significantly, providing maps free of sidelobe contamination. Nevertheless, the increased sparsity of the arrangement of the focal points in the observation grid as determined by the coordinate transformation degrades the resolution at high ranges. This can be compensated by the proposed dynamic focusing scheme. The simulations are consistent with acoustic measurements indicating the applicability of the method in practice.

Acknowledgements

The authors would like to thank Jørgen Hald and Jesper Gomes, Brüel & Kjær, for useful discussions and for providing a prototype microphone array with 60 microphones and other equipment for the experimental part of this investigation.

References

[1] D. Johnson, D. Dudgeon, *Array Signal Processing*, Prentice Hall, 1993.
[2] T. Brooks, W. Humphreys, A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays, *Journal of Sound and Vibration* 294 (2006) 856–879.
[3] R. Dougherty, Extensions of DAMAS and benefits and limitations of deconvolution in beamforming, in: *11th AIAA/CEAS Aeroacoustics Conference*, volume 3, AIAA, 2005, pp. 2036–2048.

[4] K. Ehrenfried, L. Koop, A comparison of iterative deconvolution algorithms for the mapping of acoustic sources, *AIAA Journal* 45 (2007) 1584–1595.
[5] M. Palmese, A. Trucco, Three-dimensional acoustic imaging by chirp zeta transform digital beamforming, *IEEE Transactions on Instrumentation and Measurement* 58 (2009) 2080–2086.
[6] M. Palmese, A. Trucco, Digital near field beamforming for efficient 3-D underwater acoustic image generation, in: *IEEE International Workshop on Imaging Systems and Techniques - IST 2007*, 2007, pp. 33–37.
[7] L. Ziomec, Three necessary conditions for the validity of the Fresnel phase approximation for the near-field beam pattern of an aperture, *IEEE Journal of Oceanic Engineering* 18 (1993) 73–76.
[8] G. Elias, Source localisation with a two-dimensional focused array: optimal signal processing for a cross-shaped array, in: *Proceedings of Inter-Noise*, volume 2, 1995, pp. 1175–1178.
[9] J. Christensen, J. Hald, Improvements of cross spectral beamforming, in: *32nd International Congress and Exposition on Noise Control Engineering (Inter-Noise)*, 2003.
[10] J. Hald, J. Christensen, A novel beamformer array design for noise source location from intermediate measurement distances, *Journal of the Acoustical Society of America* 112 (2002) 2448–2448.
[11] V. Murino, A.Trucco, Dynamic focusing by FFT beamforming for underwater 3D imaging, *Acoustic Letters* 17 (1994) 169–172.
[12] V.Murino, A. Trucco, Underwater 3D imaging by FFT dynamic focusing beamforming, in: *Proceedings ICIP-94*, volume 1, 1994, pp. 890–4.
[13] W. T. Vetterling, S. A. Teukolsky, W. H. Press, B. P. Flannery, *Numerical Recipies in C: The Art of Scientific Computing*, Cambridge University Press, 1992.
[14] H. H. Hubbard, K. P. Shepherd, Wind Turbine Acoustics, Technical report NASA TP-3057, DOE/NASA/20320-77, National Aeronautics and Space Administration, Langley Research Center, Hampton, VA 23665-5225, 1990.
[15] P. Moriarty, P. Migliore, Semi-Empirical Aeroacoustic Noise Prediction Code for Wind Turbines, Technical Report NREL/TP-500-34478, National Renewable Energy Laboratory, 1617 Cole Blvd. Golden,CO 80401-3393, 2003.