OPEN ACCESS Freely available online

Audio Engineering Society

Conference Paper 39

Presented at the International Conference on
Immersive and Interactive Audio
2019 March 27–29, York, UK

This paper was peer-reviewed as a complete manuscript for presentation at this conference. This paper is available in the AES E-Library (http://www.aes.org/e-lib) all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Acoustically hard 2D arrays for 3D HOA

Svein Berge

Harpex Ltd, Wilhelm-Stolze-Straße 34, 10249 Berlin, Germany

Correspondence should be addressed to Svein Berge (sveinb@harpex.net)

ABSTRACT

The acquisition of higher-order ambisonic signals presents a technical challenge which has been met by several authors proposing a number of different array geometries and sensor types. The current paper introduces a class of arrays whose performance has not been assessed before; arrays comprising only pressure-sensitive sensors on both sides of an acoustically hard plate. The combined acoustical and signal processing system is analyzed and numerical experiments on an optimized third order array provide a performance comparison with a more conventional hard-shell spherical microphone array. The model is verified by measurements.

1 Introduction

Higher-order ambisonics (HOA) has in recent years established itself as an attractive format and technology for storage, manipulation and distribution of spatial audio content. It is based on the idea that any sound field can be represented with arbitrary precision by truncating an infinite set of basis functions [1]. The ambisonic basis functions are

$$F_l^m(\theta,\phi,kr) = P_l^m(\cos\theta)\exp(im\phi)\,j_l(kr), \quad (1)$$

where P_l^m are the associated Legendre polynomials, j_l are the spherical Bessel functions, θ is the polar angle, φ is the azimuthal angle, r is the radius and k is the wave number ω/c, where ω is the angular frequency and c is the speed of sound. The indexes l and m are called the basis function's *order* and *degree,* respectively. The values for l are the nonnegative integers, and the values for m are the integers $-l \ldots l$. Different normalizations of P_l^m exist in the literature. Throughout this paper, we will assume N3D normalization [1].

In free field conditions, any sound field $V(\theta,\phi,kr,\omega)$ can be precisely expressed as the sum

$$V(\theta,\phi,kr,\omega) = \sum_{l,m} y_l^m(\omega)F_l^m(\theta,\phi,kr). \quad (2)$$

The goal of an ambisonic acquisition system is to produce a finite set of band-limited signals $y_l^m(t)$ which approximates the sound field V. Since any practical microphone array will cause some perturbation of the sound field, it should be noted that V refers to the incident field, i.e. the field which would have been observed in the absence of any scattering or other form of perturbation by the array itself.

Open spherical microphone arrays can acquire HOA signals [2][3], but a single array of pressure-sensitive microphones is generally not sufficient because of radial zero crossings in the ambisonic basis functions, leading to nulls in the frequency responses. This can be solved either by using several concentric arrays or first-order microphones.

However, the most commonly used solution is to mount the microphones on an acoustically hard shell [4]. The scattering off the shell effectively removes the nulls in the frequency response [2]. This scattering causes more perturbation of the sound field

than an open array, but the perturbation can be accurately modeled and accounted for in the signal processing.

Open, planar arrays have also been proposed as means to acquire ambisonic signals [5], but much like open spherical arrays, these arrays are also subject to nulls in the basis functions. Without loss of generality, a planar array can be defined to lie in the x-y plane, where about half of the basis functions are zero at all frequencies. Similarly to spherical arrays, the problem can be solved either by using several parallel arrays or first-order microphones.

The objective of this paper is to find out if similar advantages can be gained for planar arrays as for spherical arrays by instead introducing a scattering surface, in this case an acoustically hard plate. This could have advantages in terms of manufacturability and use. For example, if 360° video is required, a flat array could accommodate camera systems at the center of both sides that can be stitched together with minimal parallax error.

In Section 2, the physical structure and an idealized acoustical model of the array will be introduced. In Section 3, a method for producing the desired output signals from the measured signals is proposed. In Section 4, the performance of the acoustical and signal processing systems is assessed for a particular choice of array parameters, both generally and for the particular use case of beamforming. In Section 5, the idealized acoustical model is compared to measurements.

2 Physical structure

The arrays studied here consist of a circular, rigid and acoustically hard disc with omnidirectional microphones distributed over both surfaces, as illustrated in Figure 1. We define the plane of the disc to be the x-y plane.

2.1 Boundary conditions

At the surface of an acoustically hard plate, the normal component of the pressure gradient is zero. We will assume that the plate has negligible thickness, yet does not move or deform in response to the incident field.

2.2 Total sound field

Let a denote the radius of the disc and let us name the total field on the surfaces of the disc $V_{l\pm}^m$, where + and - refer to the two sides of the disc, and the incident field is described by one of the ambisonic basis functions defined in Equation 1. The functions $V_{l\pm}^m$ are separable into an azimuthal and a radial part:

$$V_{l\pm}^m(\phi, r) = \exp(im\phi) R_{l\pm}^m(r/a, ka), \quad (3)$$

$$R_{l\pm}^m(r/a, ka) = P_l^m(0) j_l(kr) \pm S_l^m(r/a, ka). \quad (4)$$

The two terms in $R_{l\pm}^m$ represent the incident and the scattered field, respectively. An analytical expression for the radial scattering functions S_l^m can presumably be derived from the results in [6], where an analytical expression for the scattering of plane waves by a circular disc is given. For shapes other than circular discs, the numerical method presented in [7] can be used to calculate the scattered field. Since non-circular shapes were considered early on in this project, that was the method used also for the case of circular discs.

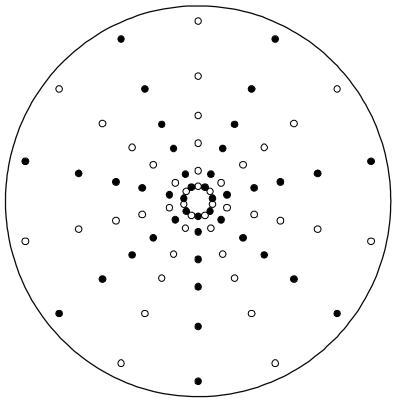

Figure 1. Example of planar array: $N_\phi = 7$, $N_r = 6$, $\alpha = 360°/28$. ○ microphones on top side ● microphones on bottom side

The basis functions can be grouped into even and odd functions, according to the parity of the number $l + m$.

A few properties of $V_{l\pm}^{m}$ can be noted:

1) Even F_l^m have no z-gradient in the x-y plane, so no scattering takes place. The total field is the same on both surfaces of the plate, and is given by F_l^m.
2) Odd F_l^m are zero in the x-y plane, so only the scattered field contributes to the microphone signals. The field is opposite on both surfaces of the plate.
3) Due to the cylindrical symmetry of the system, the azimuthal part of the scattered field is equal to that of the incident field, $\exp(im\phi)$.
4) Due to the continuity and odd symmetry of the scattered field, the scattered field is zero at the edge of the plate.

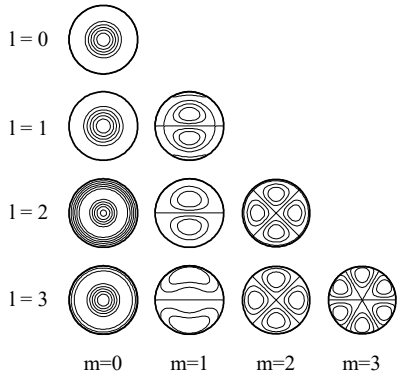

Figure 2. Surface pressure V_{l+}^{m} (real part), ka = 6.1.

2.3 Plate shape

As a consequence of this last property, the signal strength of odd modes at low frequencies increases as the distance from the edge of the plate increases. The center of a circular disc is farther removed from the nearest edge than any point on a plate of the same area with any other shape. For this reason, circular discs provide greater low frequency signal strengths than other 2D shapes.

2.4 Microphone location

In order to produce a full set of signals, it is necessary to distinguish between all the desired incident modes. Judging from Figure 2, some of the modes result in surface pressure distributions that are very similar to each other. For example, mode (0,0) and mode (1,0) look almost identical. However, since the parities of these modes are different, they are easily distinguished when combining signals from both sides of the disc due to properties 1 and 2. For this is the reason, it is necessary to place microphones on both sides of the disc.

Since the sound fields are separable into radial and angular functions, it seems reasonable to constrain the microphone placement to corresponding configurations, i.e. concentric rings of microphones.

Due to property 3, the azimuthal part of all the desired signals takes the form of sinusoids. If the ambisonic order is limited to $l \leq L$, the desired angular frequencies will be limited to the range $-L..L$. This involves $2L + 1$ different sinusoidal components, which can be sampled using a minimum of $N_\phi = 2L + 1$ uniformly spaced microphones per ring.

The radial part of the total sound field contains zero crossings, except at the lowest frequencies. The locations of the zero crossings are frequency dependent. This means that any single ring of microphones will be deaf to certain modes at certain frequencies. It is therefore necessary to use more than one ring. The number of rings and their radiuses will become an optimization problem which is described in Section 4.1.

Finally, because some microphone types by construction cannot be placed in pairs directly opposite each other, it may be necessary to displace the bottom microphones relative to the top microphones, as shown in Figure 1.

3 Ambisonic encoding

Given that the response x_i of each microphone to each of the component incident fields F_l^m is known, the calculation of the ambisonic output signals y is in principle only a matter of multiplying the observation vector x with the inverse of the array response matrix.

However, if the microphones are placed as in Figure 1, we can separate the signals by order, parity and radial response in three separate stages, as

shown in Figure 3, thus saving resources. Firstly, a Fourier transform applied to signals from the same radius on the same side of the disc resolves the signals by degree (m). This will be called the azimuthal stage. Secondly, adding and subtracting signals from the two sides of the disc separates even and odd modes. This will be called the parity stage. However, if the microphones on the two sides are not directly opposite each other, the phases of the azimuthal stage outputs will be shifted by a phase angle of $\pm\alpha m$, which must be adjusted accordingly before the parity stage.

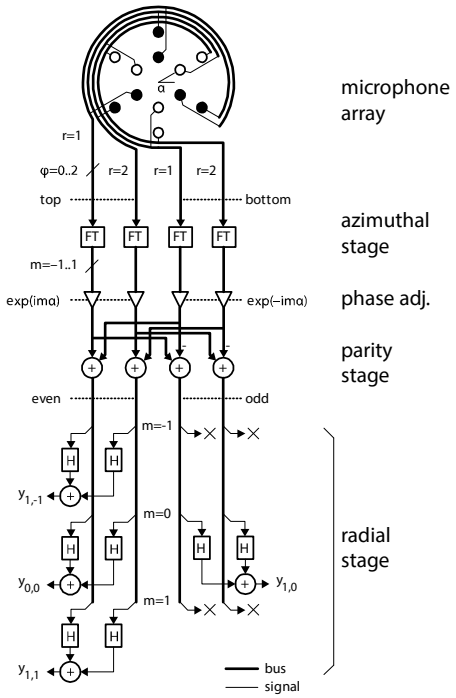

Figure 3. Encoding stages for a 3×2×2 array

Finally, in the radial stage, the contributions with equal degree and parity from different radiuses are combined using a set of filters.

Since we sample the sound field at discrete locations, there will be aliasing from orders and degrees higher than the ones we wish to sample. It is the job of the radial stage filters to suppress these.

When the microphones on both sides of the disc are directly opposite each other (eclipsed arrangement), the aliasing follows the aliasing pattern familiar from time domain sampling, shown in the top panel of Figure 4. However, when the microphones are staggered, a different pattern emerges. Some even modes will be detected with opposite polarity on the two sides of the disc, and some odd modes will be detected with equal polarity on the two sides, leading to an inversion of the apparent parity for modes of certain degrees, shown in the bottom panel of Figure 4. This needs to be taken into account when constructing the radial stage filters.

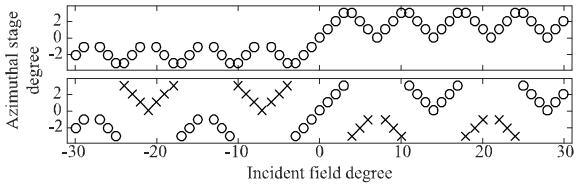

Figure 4. Alias mapping at the output of the parity stage, $N_\phi = 7$. Top: Eclipsed arrangement, Bottom: Staggered arrangement. ○: Parity is preserved ×: Parity is inverted

3.1 Radial stage filters

The output from the parity stage is grouped by degree and parity. Let us collect the signals in each group into a vector $p_\pm^m$. There will be $2N_\phi$ vectors and each vector will have N_r elements, where N_r is the number of rings in the array. Each vector will contain contributions from several modes, but each mode will only contribute to one vector, if any. Which vector is affected by which mode is given by the alias mapping. For example, for an $N_r = 7$ staggered microphone layout, it can be read from Figure 4 that the vector p_+^1 would have contributions from the desired components F_1^1 and F_3^1 along with the aliased components F_7^{-6}, F_9^{-6}, F_9^{-8}, F_{11}^{-6}, F_{11}^{-8}, F_{13}^{-6}, F_{13}^{-8}, F_{13}^{13}, F_{15}^{-6}, F_{15}^{-8}, F_{15}^{13}, F_{15}^{15}, … There are infinitely many aliased components, but they quickly decrease in strength as l increases beyond ka, allowing the sequence to be truncated. To simplify notation, let us consider only one vector of parity stage signals, p, and group the subset of coefficients y_l^m and radial functions R_{l+}^m that contribute to this vector into the vector y and the matrix $\mathbf{R}$:

$$p = \mathbf{R}y. \tag{1}$$

We need to invert this in order to find the output signals:

$$y = \mathbf{H}p + \epsilon. \quad (2)$$

We find an approximate inverse using Tikhonov regularization:

$$\mathbf{H} = (\mathbf{R}^*\mathbf{R} + \beta\mathbf{I})^{-1}\mathbf{R}^*, \quad (3)$$

where * is the conjugate transpose and β is a trade-off parameter, ranging from near 0 for the least squares solution towards infinity for the least noise solution. If we want unity gain for the non-aliased components, we need to normalize the filters:

$$\mathbf{H}' = \mathrm{D}(\mathbf{HR})^{-1}\,\mathbf{H}, \quad (4)$$

where the D operator sets all non-diagonal elements of its argument matrix to zero.

4 Numerical experiments

For our numerical experiments, we optimize a 3rd order array for use in the audible range. As a reference we also calculate the response of a hard-shell spherical array with the same number of microphones.

A 3rd order array requires at least 7 microphones per ring, so we select $N_\phi = 7$. We also need to select N_R, a and the radius of each ring of microphones. The latter is selected through the optimization procedure described in Section 4.1. Different values for a and N_R are tested. Increasing a improves low-frequency response at the cost of lowering the aliasing frequency. Increasing N_R improves the bandwidth, particularly at high orders, at the cost of a larger number of microphones. It is found that $N_R = 6$ is the minimum required for the desired frequency range of 15 kHz, and $a = 8.5$ cm provides a reasonable trade-off between low-frequency response and manufacturability.

The locations of the microphones in the spherical reference array are selected by minimizing the Coloumb potential of 84 point charges constrained to a sphere. It is not known whether this is the optimal location of microphones, but the method does result in the optimal microphone locations in the cases where that is known [8]. Spherical t-designs [9] were also tested, but yielded lower aliasing frequencies. A sphere radius of 4 cm resulted in a similar frequency range for the two arrays.

4.1 Optimization

The parameters left for optimization are the radiuses of the microphone rings, of which there are 6. There are two possible criteria to optimize for; noise and shape error.

Noise performance of microphone arrays is characterized by their white noise gain, i.e. the amount of noise attenuation relative to a single sensor. In this case, we have

$$\mathrm{WNG}_i = 1/\sqrt{\textstyle\sum_j \left|H'_{i,j}\right|^2}. \quad (5)$$

The shape error is the magnitude of the difference between the ideal and the actual polar patterns, and includes the effects of aliasing (components with $l > 3$) and cross-talk (wrong components with $l \leq 3$). The shape error of a given mode is defined as the l^2 norm of the non-diagonal elements of the corresponding row of $\mathbf{H}'\mathbf{R}$, and is evaluated for orders up to 35. Higher orders make a negligible contribution in the studied frequency range.

For the optimization, we choose β so as to constrain the shape error to a maximum of -12 dB across the frequency range for every output channel and select the radiuses that minimize the A-weighted noise under this constraint. We assume that the microphones have a white noise characteristic with an A-weighted magnitude of 0 dB and optimize over a frequency range up to 15 kHz, starting at 0 Hz, 0 Hz, 242 Hz and 819 Hz for the orders 0-3, respectively. These are the frequencies where the WNG for the spherical array crosses -30 dB.

The Nelder-Mead method [10] is used for the optimization, which converges to the solution given in Table 1.

Ring no.	1	2	3	4	5	6
Radius / mm	8.1	18.1	27.9	39.6	55.8	78.6

Table 1. Model optimization result.

4.2 Overall performance

The performance of the resulting array is shown in Figure 5, along with the performance of the spherical array. The WNG and shape error is calculated for each mode individually.

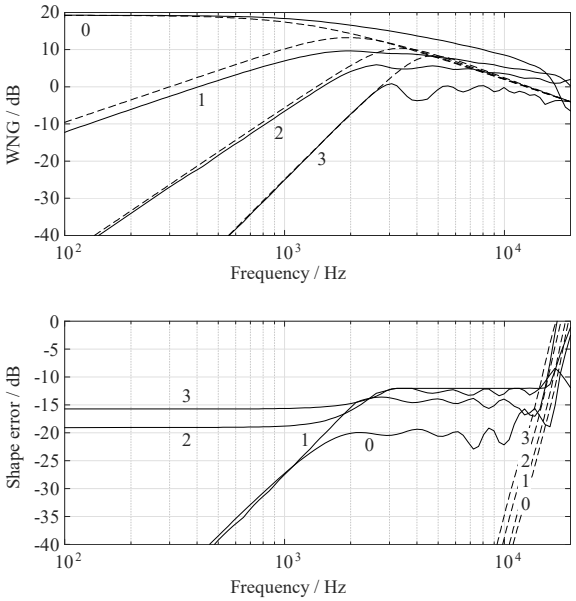

Figure 5. Average white noise gain (higher is better) and shape error (lower is better) for each ambisonic order for the spherical array (– –) and the planar array (—).

Unlike the spherical array, each degree within each order has a slightly different WNG and shape error in the planar array. For clarity, only the root mean square values for each order are plotted.

The radiuses of the two arrays were chosen to produce similar asymptotical WNG at low frequencies. In the planar array, the even modes have slightly higher WNG than the odd modes (not shown).

At 2–15 kHz, the planar array has an overall flatter WNG than the spherical array. This can be understood as a consequence of the multi-radius nature of the array. As the frequency increases, the radial stage filters will shift the weight from the outer microphones to the inner ones, which are better located to pick up high-frequency signals. However, at any given frequency there are fewer microphones in effective use, which explains why the WNG peaks at a lower value for the planar array than for the spherical one, except for the 0th order component.

There is some waviness in the WNG of the planar array, particularly at 3rd order. As the radial zero crossings of $V_{l\pm}^{m}$ move across the disc, the radial stage filters need to shift the weight around to avoid them, which causes this effect.

In any ambisonic array, the bandwidth is limited by aliasing at high frequencies and noise at low frequencies. In our example, these aspects of the two arrays are similar. There are differences in the mid-frequency noise level. However, if we assume that the recordings are intended for human listeners, we must apply a suitable weighting curve and integrate the noise level across the spectrum in order to predict its importance. Table 2 summarizes the result numerically as the A-weighted noise across the optimized part of the spectrum. The two arrays produce noise levels within 1 dB of each other for orders 2 and 3, since most of the noise comes from the low-frequency region, where both arrays perform similarly. The planar array performs better for orders 0 and 1.

Order	0	1	2	3
Cutoff	0 Hz	0 Hz	242 Hz	819 Hz
Spherical array	-4.9 dB	-4.8 dB	3.5 dB	11.6 dB
Planar array	-10.8 dB	-6.7 dB	4.3 dB	12.0 dB

Table 2. Model noise level (lower is better), A-weighted

4.3 Beamforming

An ambisonic signal is always decoded before its final use, and we would ideally like to know the effect of the shape errors on the decoded result for every conceivable type of decoding. To this end, we will study their effect on the simplest possible decoding; a single-direction beamformer. From this it should be possible to extrapolate their effect also on more complex decoding schemes.

It should be noted that if beamforming is the ultimate objective, any array will perform better if the beamformer is given access directly to the microphone signals rather than going through an ambisonic encoder [11]. This would also affect the optimization criterion, and the results of such a process are beyond the scope of the current paper.

Next to its noise figure, the most important performance characteristic of a beamformer is its directivity index, DI. This is the ratio of its sensitivity in the steering direction to its average sensitivity in all directions. The highest theoretical DI attainable with an ambisonic signal is $(N+1)^2$, where N is the ambisonic order. For a 3rd order signal, this ratio is 16, or about 12.0 dB.

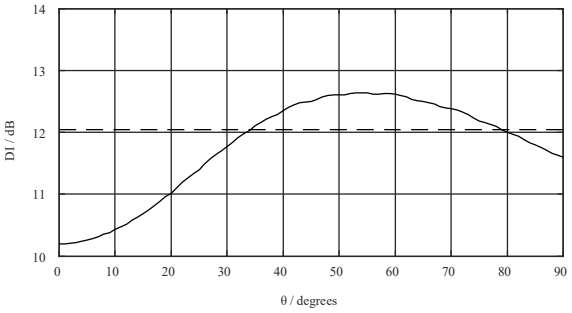

Figure 6. Directivity index as a function of steering direction for an ideal 3rd order signal (– –) and the planar array (—) at f = 5 kHz.

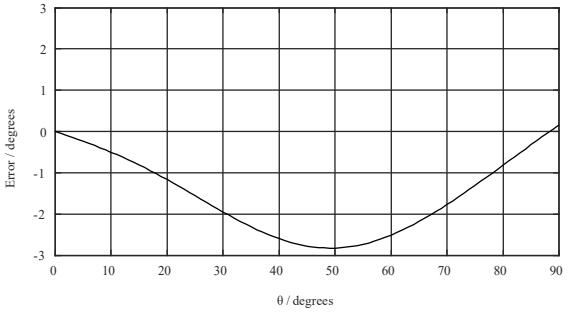

Figure 7. Error in steering direction for the planar array at f = 5 kHz.

Another important quantity is the error in steering direction. For a given directivity response $\psi(\vec{r})$ defined on the unit sphere, we can calculate an effective steering vector by integrating over the sphere:

$$\vec{s} = \int_\Omega \vec{r}\, |\psi(\vec{r})|^2 \, d\vec{r}. \tag{6}$$

The angle between the expected steering direction and the effective steering vector is the steering direction error.

These two quantities are plotted in Figures 6 and 7 for one particular frequency. The mean error is 1.8° and the mean DI is 12.1 dB, both integrated over the sphere (i.e. weighted by $\sin\theta$). Figure 8 shows the beam shapes in the two directions where they would be expected to be maximally deformed, since the value of DI deviates maximally from 12 dB in these directions.

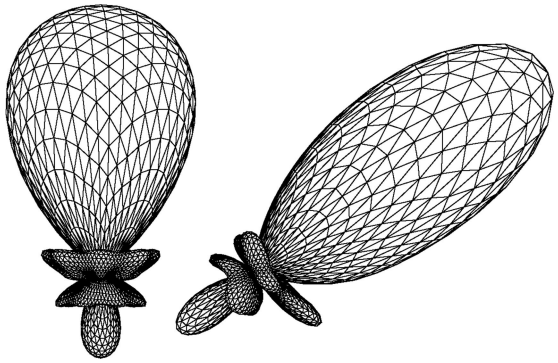

Figure 8. Beam shapes at f = 5 kHz. Left: $\theta = 0°$. Right: $\theta = 53°$. The radial component in these polar plots represents linear amplitude using the same scale for both beams.

5 Experimental verification

Several experimental measurement devices were built, using an outer radius of 8.5 cm (see Figure 9). The noise profile of the microphones was measured and incorporated into the optimization process and the ITU-R 468 noise weighting was used, resulting in the radiuses in Table 3.

Ring no.	1	2	3	4	5	6
Radius / mm	6.7	13.1	25.3	37.1	54.2	78.3

Table 3. Device optimization result.

Each device contains 84 Knowles SPH0641 MEMS microphones, an FPGA and a USB interface, all mounted on a circular, 0.6 mm thick circuit board made of the composite material FR-4. The device also contains a thermometer circuit and an ultrasonic piezoelectric transducer, which are used to measure the local speed of sound either directly or indirectly, since the radial filters need to be adjusted accordingly.

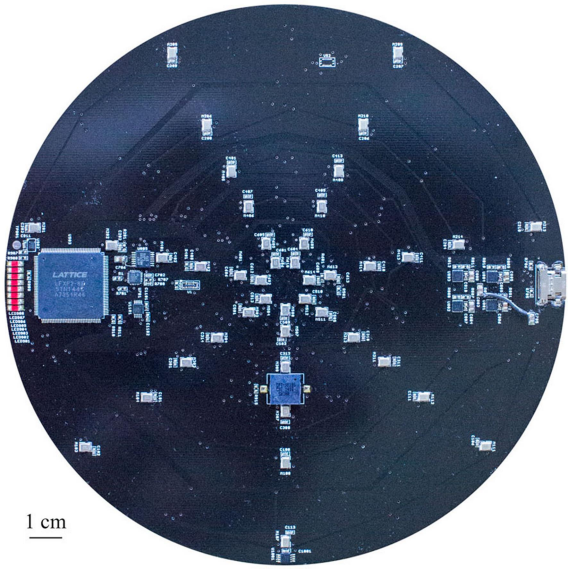

Figure 9. Experimental measurement device

5.1 Measurement setup

The device is suspended from the ceiling using a 3 mm brass tube with a length of 1.5 m (Figure 10). Power and signals are sent through wires inside this tube. One end of the tube is connected to the edge of the circuit board and the other end is attached to an angle gauge, allowing measurements to be taken at a series of rotations about the device's x axis. The tube is stabilized with guy wires to prevent lateral movement of the device during rotation.

A loudspeaker is placed 2 m away from the device. The loudspeaker consists of two concentric drivers with a crossover frequency of around 5 kHz and is housed in an airtight, axisymmetric enclosure. Apart from the loudspeaker, the device and their supports, there are no objects or structures within a volume with less than 1 m additional path length. The impulse response measurements should therefore be free from external reflections up to 2.9 ms, and only the first 1.5 ms are used in the following. The impulse responses are measured according to the methods in [12] for every 5° of θ from -90° to 90°.

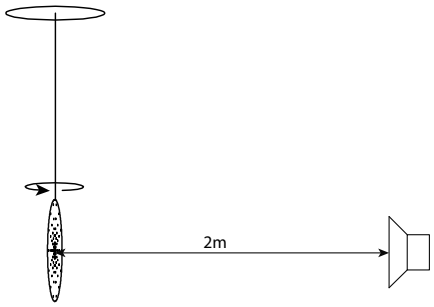

Figure 10. Experimental setup (figure not to scale)

5.2 Processing

The impulse responses are equalized using the sum of the signals from the inner ring of microphones on both sides of the board, boosted by a factor of $1/j_0(kr_1)$, as a reference. This removes the frequency response characteristics of the loudspeaker and microphones from the data.

The sensitivity of the microphones varied with a standard deviation of 0.24 dB. Compensating for these differences had no discernible effect on the results.

The measured data is compared with the model, using orders up to 35, under the assumption that the sound source produces plane waves. The measured impulse responses are sent through the azimuthal and parity stages of the encoder. This results in 48 image pairs of the type shown in Figure 11. Some of the parity stage outputs show systematic differences between the model and measurements. These differences may be related to aspects of the measurement device or the experimental setup that are not captured by the model. The best solution would be to refine the model and/or the experimental setup so as to eliminate systematic differences. Until that option is available, we would like to quantify the effect of these differences on the performance of the array.

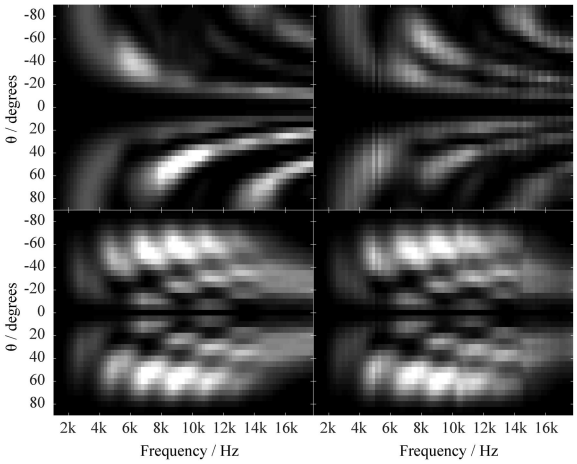

Figure 11. Comparison between model (left) and measured data (right) at two outputs from the parity stage. Top: Even parity, $m = 2$, $r = 5$. Bottom: Odd parity, $m = 1$, $r = 2$.

One way of doing this is to calculate new radial stage filters based on the measurements, using the same shape-error constraints as was used for the model-based filters. Since the device has been optimized using an incomplete model, we expect the WNG of the measurement-based filters to be lower than that of the model-based filters. The magnitude of the difference represents the potential improvement that a refined model could bring.

5.3 Results

For the measurement-based filters, we again use Tikhonov regularization and choose the tradeoff parameter such as to limit the shape error to a maximum of -12 dB for each channel. The WNG of these filters are compared to those of the model-based filters.

Up to about 12 kHz, the two WNG curves in Figure 12 are very close, and the deviation remains moderate up to about 15 kHz. Only above this frequency do the measurement-based filters underperform, leading to the conclusion that the systematic errors are not significant for the assessment or optimization of the performance of the array. A slight ripple around 5 kHz coincides with a similar artifact in the output from the parity stage as seen in Figure 11. This might be due to phase mismatch between the two loudspeaker drivers around the crossover frequency.

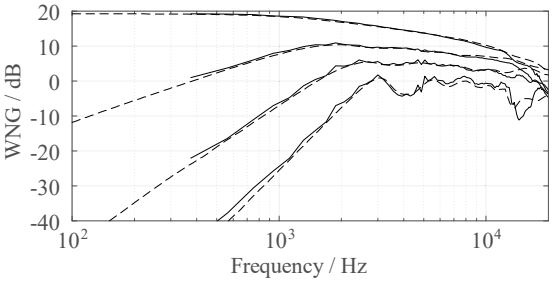

Figure 12. Average white noise gain for each ambisonic order based on the modeled (– –) and the measured (—) response

6 Conclusion

A planar microphone array consisting only of omnidirectional microphones distributed over both surfaces of an acoustically hard disc is capable of acquiring a full set of ambisonic components. The resulting scattering problem is mathematically harder to solve than that of a spherical array, but it is not intractable. A simplified acoustical model of the system captures its function to a degree of accuracy sufficient to predict its performance and optimize its design parameters.

Assessing the performance of this class of array requires detailed analysis. Using hard-shell spherical arrays as a reference, the major differences are that the proposed arrays are strictly limited to the ambisonic order they are designed for and that they introduce a vertical directional error, albeit below the threshold of perception [13].

In the studied example, the design order was 3, and the disc radius was 2.1 times the radius of the spherical reference array. This resulted in the same frequency range for the two arrays and 3rd order A-weighted noise levels within 0.4 dB of each other.

In applications requiring a fixed number of broadband channels, practical considerations relating to manufacturing and use will probably outweigh these performance differences.

7 References

[1] Daniel, Jérôme. "Acoustic field representation, application to the transmission and the reproduction of complex sound environments in a multimedia context." Ph. D. Thesis, Universite Paris (2000).

[2] Rafaely, Boaz. "Analysis and design of spherical microphone arrays." *IEEE Transactions on speech and audio processing* 13.1 (2005): 135-143.

[3] Abhayapala, Thushara D., and Darren B. Ward. "Theory and design of high order sound field microphones using spherical microphone array." *ICASSP*. Vol. 2. 2002.

[4] Meyer, Jens, and Gary Elko. "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield." *Acoustics, Speech, and Signal Processing (ICASSP),* 2002 IEEE International Conference on. Vol. 2. IEEE, 2002.

[5] Chen, Hanchi, Thushara D. Abhayapala, and Wen Zhang. "Theory and design of compact hybrid microphone arrays on two-dimensional planes for three-dimensional soundfield analysis." *The Journal of the Acoustical Society of America* 138.5 (2015): 3081-3092.

[6] Bowman, John J., Thomas B. Senior, and Piergiorgio L. Uslenghi. “The disc.” *Electromagnetic and acoustic scattering by simple shapes*. Michigan Univ Ann Arbor Radiation Lab, 1970.

[7] Williams, Earl G. "Numerical evaluation of the radiation from unbaffled, finite plates using the FFT." *The Journal of the Acoustical Society of America* 74.1 (1983): 343-347.

[8] Elko, Gary, Kubli, Robert, Meyer, Jens, US 7587054 B2, United States Patent and Trademark Office, US 7,587,054, 8 September 2009.

[9] Hardin, Ronald H., and Neil JA Sloane. "McLaren’s improved snub cube and other new spherical designs in three dimensions." Discrete & Computational Geometry 15.4 (1996): 429-441.

[10] Nelder, John A., and Roger Mead. "A simplex method for function minimization." The computer journal 7.4 (1965): 308-313.

[11] Rafaely, Boaz, et al. "Spherical microphone array beamforming." *Speech Processing in Modern Communication*. Springer, Berlin, Heidelberg, 2010. 281-305.

[12] Farina, Angelo. "Simultaneous measurement of impulse response and distortion with a swept-sine technique." *Audio Engineering Society Convention 108. Audio Engineering Society,* 2000.

[13] Carlile, Simon, Philip Leong, and Stephanie Hyams. "The nature and distribution of errors in sound localization by human listeners." *Hearing research* 114.1-2 (1997): 179-196.