BEAMFORMING WITH DOUBLE-SIDED, ACUSTICALLY HARD PLANAR ARRAYS

Svein Berge
Harpex Ltd
sveinb@harpex.net

ABSTRACT

A disc-shaped baffle with an array of microphones on each surface has recently been proposed as a device for acquiring higher-order ambisonic signals. In this paper, we will study how an array of this type performs with regards to three different beamforming algorithms. The results are quantified through numerical experiments and verified by measurement.

1. INTRODUCTION

Numerous geometries and sensor types have been studied for the purpose of producing microphone arrays with good beamforming properties while at the same time conforming to various practical and economic constraints.

The array studied in this paper is a double-sided array of microphones arranged on a rigid, disc-shaped baffle. There are several motivations for choosing this geometry. Such arrays could be produced at low cost using normal electronics manufacturing techniques. They would be both compact and mechanically robust. They would take up a minimal amount of space in 360 degree video recordings, and could even vanish completely if combined with two half-sphere camera system placed at its center. However, this all is only useful if they also exhibit good acoustical properties.

Arrays of this type have recently been studied for the purpose of acquiring higher-order ambisonic signals [1]. When compared to the more conventional spherical geometry, they have both advantages and disadvantages. The disadvantages stem from the fact that the array has a different symmetry than the desired basis functions, requiring more complex encoding filters that have a lower peak white noise gain, WNG. The advantages stem from the multi-radius nature of these arrays, leading to a wider frequency range. When the resulting noise level is weighted and integrated over the spectrum, the flat array comes out on top for low ambisonic orders and ties with a comparable spherical array at orders 2 and 3.

In this paper, the beamforming performance of the flat array will be studied, following broadly the outline and notation of [2], which studied the same for spherical arrays. That article introduced several beamforming algorithms, some based on ambisonic signals derived from spherical arrays and others using the direct output of the microphone array. Here, we will only study the latter ones, since these have the highest performance.

2. ACOUSTICAL MODEL

We will model the array as a rigid, circular disc-shaped baffle. According to [3], for a plane wave incident at an angle θ_0 with the positive z-axis, such that

$$p_i = \exp\{ik(x\sin\theta_0 + z\cos\theta_0)\}, \quad (1)$$

where k is the wave number, the total field on the top surface of the disc is

$$p_i + p_s = \frac{2}{c}\sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\epsilon_m \frac{i^n}{\widetilde{N}_n^m}\frac{1}{R_{nm}^{3'}(-ika,0)}\times$$
$$S_n^{\mathrm{m}}(-ika,\cos\theta_0)S_n^{\mathrm{m}}(-ika,r/a)\cos m\phi, \quad (2)$$

where p_s is the scattered field, a is the radius of the disc, r is the distance from the center of the disc and $\widetilde{N}$, R and S are defined in [4].

For an incident field equal to one of the eigenfunctions F_n^m of the wave equation in spherical coordinates,

$$p_i = F_n^m = Y_n^m(\theta,\phi)\, j_n(kr), \quad (3)$$

where Y_n^m are the spherical harmonic functions [5] and j_n are the spherical Bessel functions, the scattered field on the top surface simplifies to

$$p_s = C_n^m S_n^m(-ika,r/a)\exp(im\phi), \quad (4)$$

where C_n^m are constants.

The incident field on the bottom surface is equal to that on the top surface, and the scattered field on the bottom surface is opposite that on the top surface.

We express a general incident field as a linear combination of eigenfunctions of the wave equation:

$$p_i = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} a_{nm}(k)Y_n^m(\theta,\phi)\, j_n(kr) \quad (5)$$

where a_{nm} are the coefficients that describe the field. The total field, as sensed by the microphones, is equal to this incident field plus the scattered field

$$p_s = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} a_{nm}(k)\, C_n^m \times S_n^m(-ika, r/a) \exp(im\phi). \quad (6)$$

The outer sums in (5) and (6) can be truncated at a finite $N > ka$, since the Bessel functions and the scattered field decrease rapidly in magnitude with n when $n > ka$. The functions are evaluated at the M microphone locations, given by r_i, ϕ_i and $\theta = 0$, bearing in mind that p_s must be negated for microphones on the bottom side of the disc. The result of all this is expressed in this matrix equation, following the notation of [2]:

$$\mathbf{p} = \mathbf{B}\mathbf{a}_{\mathbf{nm}}, \quad (7)$$

where $\boldsymbol{p} = [p(k, r, \theta_1, \phi_1), \cdots, p(k, r, \theta_M, \phi_M)]^T$ is a column vector of length M holding the pressures sampled by the array microphones and the $M \times (N+1)^2$ matrix $\mathbf{B}$ encodes the response of the array to an arbitrary sound field, where each column contains its response to an incident field equal to a single eigenfunction F_n^m. The columns are ordered by n, then by m. The vector $\mathbf{a_{nm}}$ contains the coefficients $a_{nm}(k)$, ordered in the same manner.

3. BEAMFORMERS

Of the beamformers proposed in [2], we will study the "space domain maximum directivity index with optimal alias cancellation" (SMDAC) and the "space domain maximum white noise gain with optimal alias cancellation" (SMGAC) beamformers. We will also study the "sensitivity constrained optimal beamformer" (SCOB) proposed in the context of line arrays in [6]. Apart from a normalization constant, the same beamformers can also be derived from an MVDR formulation [7].

In each case, the beamformer is expressed as a vector of weights $\mathbf{w}$ to be applied to the input signal vector $\mathbf{p}$ in order to produce the output signal y:

$$y(\mathrm{k}) = \mathbf{w}^H \mathbf{p} = \mathbf{w}^H \mathbf{B}\mathbf{a}_{\mathbf{nm}}. \quad (8)$$

Its response A in a given direction (θ_0, ϕ_0) can be found by setting $\mathbf{a_{nm}}$ equal to the coefficients of a planewave, $\mathbf{Y}_0^* = [Y_0^0(\theta_0, \phi_0), \cdots, Y_N^N(\theta_0, \phi_0)]^H$:

$$A(k, \theta_0, \phi_0) = \mathbf{w}^H \mathbf{B}\mathbf{Y}_0^*. \quad (9)$$

3.1 Maximum-directivity beamformer

This beamformer aims to maximize the directivity factor DF of the array, i.e. the output signal power for signals coming from the look direction (θ_l, ϕ_l) relative to the average power for all possible directions of incidence. The expression for the DF is given by

$$DF = 4\pi \frac{\mathbf{w}^H \mathbf{B}\mathbf{Y}_l^* \mathbf{Y}_l^T \mathbf{B}^H \mathbf{w}}{\mathbf{w}^H \mathbf{B}\mathbf{B}^H \mathbf{w}}, \quad (10)$$

The directivity index DI is the directivity factor expressed in dB. The beamformer weights is given by

$$\mathbf{w}^{\mathrm{SMDAC}} = (\mathbf{B}\mathbf{B}^H)^{-1}\mathbf{B}\mathbf{Y}_l^*. \quad (11)$$

It is not normalized (i.e. distortion-free in the MVDR sense), but can be normalized by dividing by its response in the look direction, $A(k, \theta_l, \phi_l)$.

3.2 Maximum white noise gain beamformer

The white noise gain of a beamformer is defined as the improvement in the signal-to-noise ratio in the beamformer output relative to a single sensor in free-field conditions. Its mathematical expression in this context is

$$WNG = \frac{|4\pi A(k, \theta_l, \phi_l)|^2}{\mathbf{w}^H \mathbf{w}}. \quad (12)$$

Its maximum value is achieved with the SMGAC beamformer, defined by

$$\mathbf{w}^{\mathrm{SMGAC}} = \mathbf{B}\mathbf{Y}_l^*. \quad (13)$$

3.3 Sensitivity constrained beamformer

As we will see in the following sections, the SMDAC beamformer cannot be used across the entire spectrum in practice, due to its tendency to amplify sensor noise. However, we may still want a higher directivity than that offered by the SMGAC beamformer, so a compromise between the two might be useful. The sensitivity constrained beamformer provides this through a tradeoff parameter β:

$$\mathbf{w}^{\mathrm{SCOB}} = (\mathbf{B}\mathbf{B}^H + \beta\mathbf{I})^{-1}\mathbf{B}\mathbf{Y}_l^*. \quad (14)$$

For a given WNG, this is the beamformer which optimizes the DI [6]. Or, conversely, for a given DI it optimizes WNG. Setting $\beta = 0$ gives the highest directivity and identical weights to the SMDAC beamformer. Increasing β towards infinity gives the highest WNG. Apart from the constant factor β, which vanishes when the beamformers are normalized, this gives the same weights as the SMGAC beamformer. The optimal value for β will in practice depend on the ratio between sensor noise and ambient noise.

4. NUMERICAL EXPERIMENTS

The array studied here consists of 84 microphones, with 42 placed on either side of the disc as shown in Figure 1.

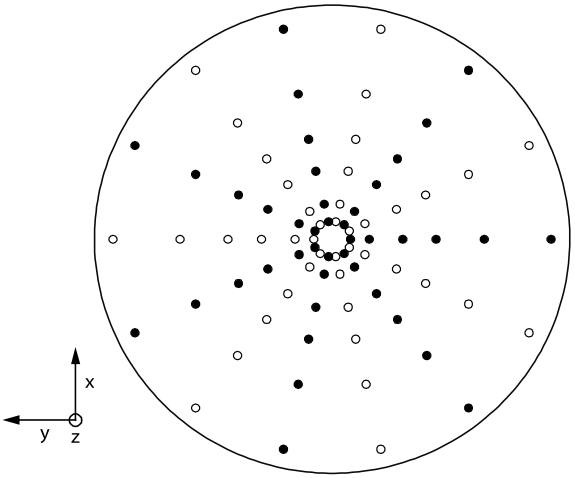

Figure 1. Microphone layout: ○ microphones on top side. ● microphones on bottom side

The layout was optimized for producing a 3rd order ambisonic signal with controlled aliasing up to a frequency of 15 kHz [1]. The array has a radius of 85 mm.

The microphones are arranged in six rings of 14 microphones, 7 on each side of the baffle. The radiuses of the rings are given in Table 1.

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

Table 1. Microphone ring radiuses.

When used far below the aliasing frequency, the SMDAC beamformer is very sensitive to noise, numerical stability and systematic errors. This makes it unsuitable for use in this frequency range, but it provides a useful upper bound for the directivity index. The SMGAC has no stability problems, but is usually not the optimal choice, since a small reduction in WNG relative to this maximum can usually provide a large increase in directivity index. As a representative of the continuum of beamformers between these two extremes we use the SCOB beamformer with $\beta = 10^{-2}$, which provides a reasonable trade-off between directivity and noise.

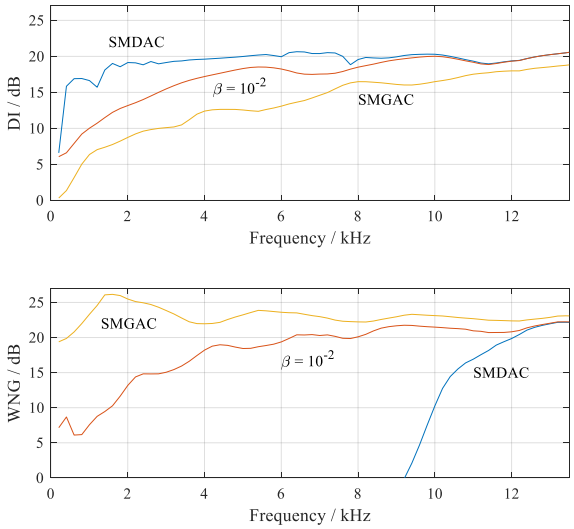

Figure 2. Directivity index and white noise gain for the three beamformers.

Figure 2 shows the directivity index and white noise gain across the spectrum for the three different beamformers when pointed in the $\theta = 0$ direction. At higher frequencies, the differences between the beamformers vanish.

Around 1.5 kHz, the scattering and SMGAC array processing combine favorably to give a maximum WNG of 26 dB, well above the maximum of $10 \log_{10} 84 = 19$ dB for an open array with the same number of microphones. Below about 3 kHz, the SMDAC plot is not reliable due to numerical instability.

Figure 3 shows the beam shapes at one frequency and illustrates how the increase in WNG comes at the cost of a wider main lobe as well as stronger side lobes. As the frequency increases, the number of side lobes will also increase, but their total energy tends to decrease. The slight increase in beam width from SMDAC to, $\beta = 10^{-2}$ provides a dramatic increase in WNG, from -44 dB to 18 dB.

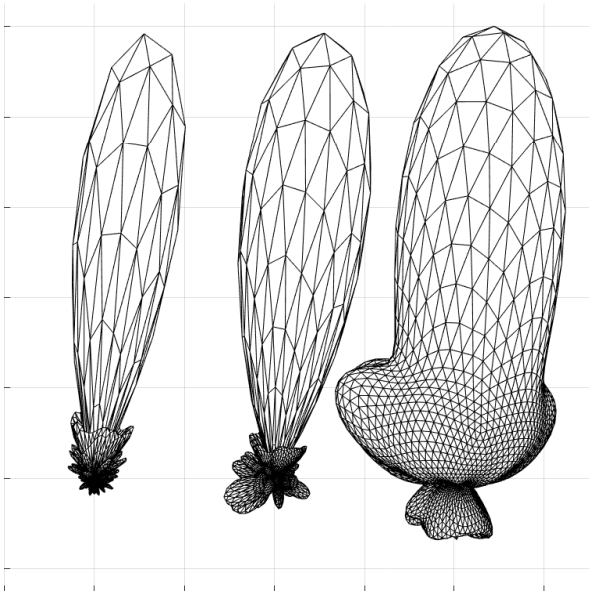

Figure 3. Beam shapes at f = 5 kHz, $\theta_l = 10°$ for the SMDAC, $\beta = 10^{-2}$ SCOB and SMGAC beamformers (left to right). The radial axes in these plots represent linear magnitude.

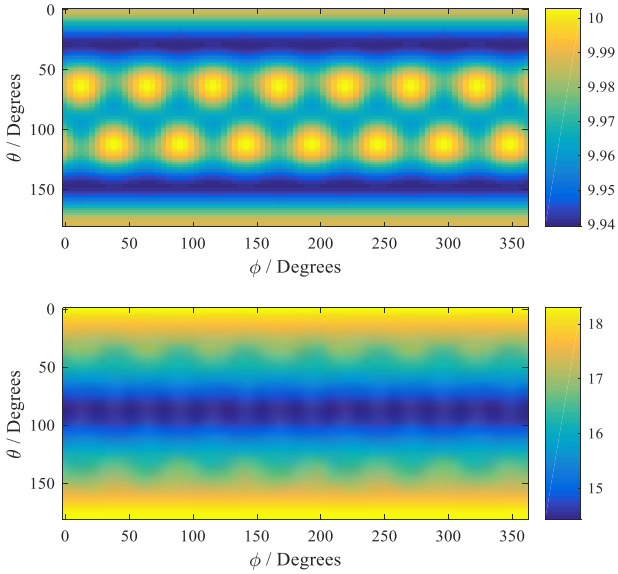

Figure 4. Directivity index for different look directions at f = 1 kHz (top) and f = 5 kHz (bottom). $\beta = 10^{-2}$ SCOB.

As one might expect from an array which is not spherically symmetrical, its directivity pattern is also not symmetrical. At low frequencies, the directivity is isotropic for all practical purposes. Above about 1 kHz, the directivity is highest along the z axis, as seen in Figure 4**.**

5. EXPERIMENTAL VERIFICATION

The correctness and practical applicability of the theoretical results is verified using a physical device consisting of 84 IM69D130 microphones placed on a 1.6 mm thick printed circuit board made from the fiberglass-based laminate FR-4 (Figure 5). The circuit board is further laminated between a 1.0 mm sheet of pressboard, a 0.3 mm polystyrene foil and 0.5 mm polyester fabric on either side. The total thickness is 5.5 mm, and the outer radius is 85 mm. The microphones are placed according to the model in the previous section. The microphones are connected to a computer via a USB interface.

Figure 5. Experimental measurement device

The device is suspended from the ceiling using a 3 mm brass tube with a length of 1.5 m (Figure 6). 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 which were driven separately and combined in post-processing with a crossover frequency of around 8 kHz. The loudspeaker enclosure is axisymmetric and airtight. The room is not anechoic. 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 [8] for every 5° of θ from -90° to 90°.

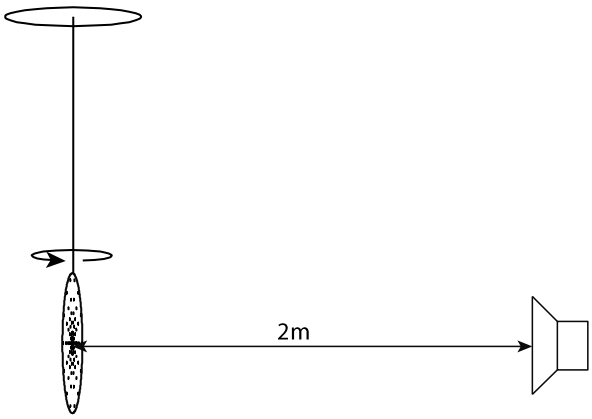

Figure 6. Measurement setup (figure not to scale).

The noise spectrum of the microphones was measured by recording the output of one microphone with a blocked acoustic port. The absolute level of the spectrum was shifted to match the A-weighted noise level of 25 dB (A) given in the device's data sheet [9]. Using this, it was possible to calculate the A-weighted equivalent noise levels of the three beamformers shown in Table 2.

Beamformer	SMDAC	$\beta = 10^{-2}$	SMGAC
Equiv. self noise	113 dB (A)	13 dB (A)	3 dB (A)

Table 2. Noise level, $\theta = 0°$.

Since the measurement setup only allows rotation about the array's x axis, we can only directly measure the beam patterns in the y-z plane, as in Figure 8. To access the beam pattern in the horizontal plane, we measure the response of the array in one horizontal direction and calculate the response as the look direction of the beamformer is rotated around the horizon, resulting in Figure 7.

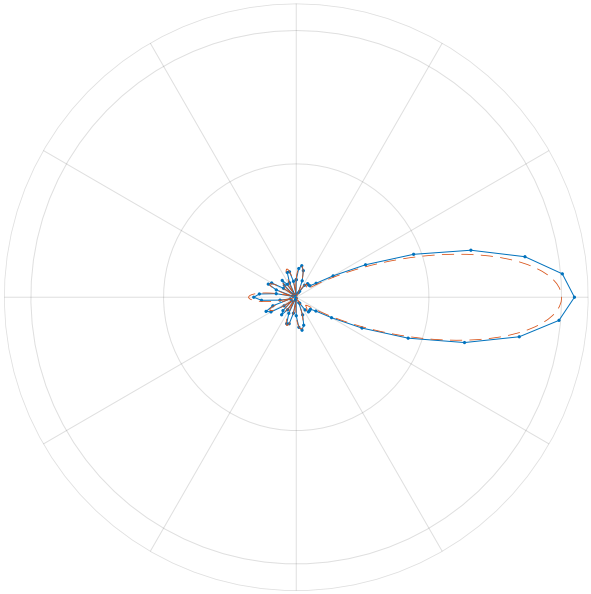

Figure 7. Modeled (– –) and measured (—) response, SCOB $\beta = 10^{-2}$ in the horizontal plane at f = 5 kHz. Radius is linear magnitude response.

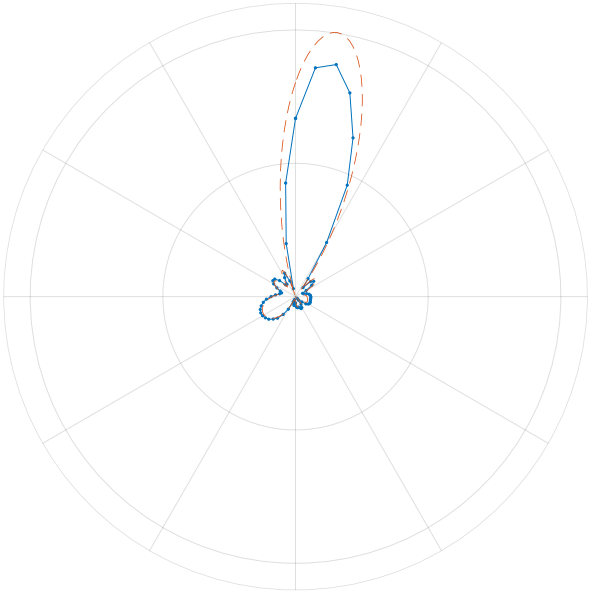

Figure 8. Modeled (– –) and measured (—) response in the y-z plane at f = 5 kHz, $\beta = 10^{-2}$ and $\theta_l = 10°$. Radius is linear magnitude response.

The same analysis that produced Figure 8 is repeated across the frequency range to produce an overview of the frequency dependency of the polar patterns. The result is shown in Figures 9 and 10. In this last figure, the front response is normalized. The measured data for the lowest frequencies (< 1 kHz) may not be reliable due to the truncated impulse response measurement method.

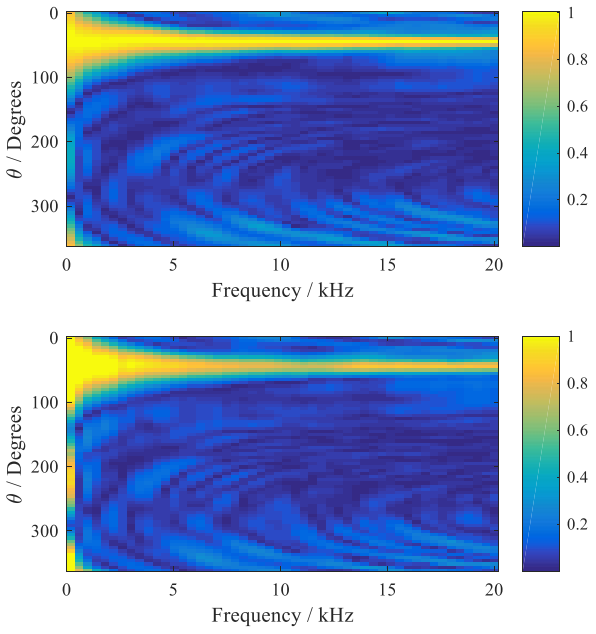

Figure 9. Modeled (top) and measured (bottom) response in a vertical plane at different frequencies, $\beta = 10^{-2}$ and $\theta_l = 45°$.

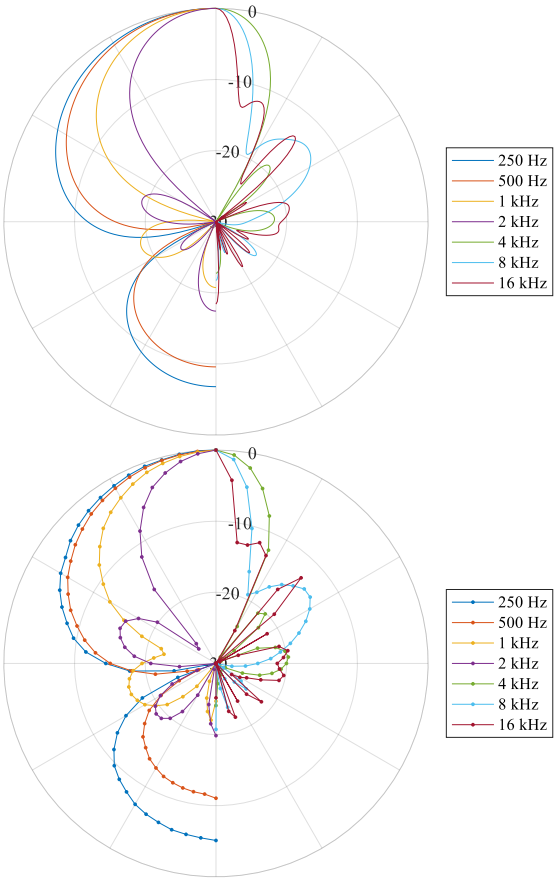

Figure 10. Modeled (top) and measured (bottom) response in the y-z plane, $\beta = 10^{-2}$ and $\theta_l = 0°$. Radius is dB magnitude response.

6. CONCLUSION

Some of the beamforming techniques that were developed for spherical arrays and line arrays also work well with double-sided disc arrays. The array was originally developed with 3rd order ambisonics in mind. However, if the ultimate goal is to perform beamforming, the direct approach studied here gives significantly better results. Array-agnostic 3rd order ambisonic beamformers are limited to 12 dB directivity index and, for arrays of this type, about 0 dB white noise gain, while the SCOB beamformer can provide a WNG of 15-20 dB with the same directivity index. It was already known that the SMDAC beamformer works best above the array's aliasing limit. Because of the multi-radius nature of the arrays studied here, the transition between no aliasing and full aliasing takes place over a much wider range than for a spherical array. The SMDAC beamformer can only be used in the upper part of this range and above.

Since double-sided disc arrays have a non-isotropic scattering function, they work better in some directions than in others. The beamformers described here take optimal advantage of any scattering that takes place. Particularly at medium to high frequencies, this effect provides a higher directivity and / or white noise gain along the z axis than in other directions. This means that for applica-

tions where the approximate direction of arrival can be predicted before setting up the microphone, this may be a better option than a spherical array, whereas in applications where an isotropic response is required, a spherical array should be used.

7. REFERENCES

[1] Svein Berge, "Acoustically Hard 2D Arrays for 3D HOA," in *Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio*, 2019.

[2] David Lou Alon and Boaz Rafaely, "Beamforming with Optimal Aliasing Cancellation in Spherical Microphone Arrays," *IEEE/ACM trans. on audio, speech, and language processing*, vol. 24, no. 1, January 2016.

[3] Frederick B. Sleator, "The Disc," in *Electromagetic and acoustic scattering by simple shapes*.: Michigan Univ Ann Arbor Radiation Lab, 1970.

[4] Carson Flammer, *Spheroidal wave functions*.: Stanford University Press, 1957.

[5] George B. Arfken and Hans J. Weber, *Mathematical Methods For Physicists*, 5th ed.: Academic, 2001.

[6] Henry Cox, Robert M. Zesking, and Theo Kooij, "Practical Supergain," *IEEE Trans. on Acoustics, speech, and signal processing*, vol. ASSP-34, no. 3, 1986.

[7] Shefeng Yan, Haohai Sun, U. Peter Svensson, Xiaochuan Ma, and Jens M. Hovem, "Optimal modal beamforming for spherical microphone arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 2, 2011.

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

[9] Infineon, "IM69D130 - High performance digital XENSIV MEMS microphone," Datasheet 2017.