Higher order loudspeakers and active compensation for improved 2D sound field reproduction in rooms

M. A. Poletti AES Member, T. Betlehem,

Callaghan Innovation, Lower Hutt, New Zealand

T. D. Abhayapala,

Australian National University, Canberra ACT 0200, Australia

Abstract

A prototype 2D surround sound system was developed to investigate the use of higher order loudspeakers and active compensation to both eliminate unwanted reflections and to improve spatial accuracy in semi-reverberant rooms. The loudspeakers, and the calibration microphone, were cylindrical designs as these are lower cost than spherical designs and better suited to 2D reproduction. This paper presents the cylindrical coordinate based theory for the calibration of the system, a description of the transducers and processing system, and experimental results obtained using a sound reproduction system consisting of five third-order loudspeakers. Approximately anechoic reproduction is achieved over a radius exceeding 90 mm and over a frequency range up to 8 kHz.

0 Introduction

Surround sound systems aim to reproduce a desired sound field within an array of loudspeakers [1]. These systems typically aim to reproduce an exact sound field over a finite region. An alternative is the perceptual approach in which only those cues which are important for creating a correct perception of the desired field are recorded and reproduced. [2]–[4].

The theory of exact sound field reproduction is typically based on the discrete approximation of the Kirchhoff Helmholtz integral equation, leading to the Wave Field Synthesis approach [5]–[7], or on cylindrical and spherical harmonic solutions to the wave equation, which lead to the higher order Ambisonics (HOA) approach [8]–[11]. The cylindrical description is in its general form applicable to the 3D case [12][13], but is typically simplified to sound fields reproduced in the horizontal plane [10].

There are two main limitations which affect the ability of surround systems to reproduce a desired sound field [14]. The use of a finite array of loudspeakers means that sound field reproduction is only possible up to a finite frequency governed by the spacing between the loudspeakers. When the loudspeakers are more than half a wavelength apart, spatial aliasing can occur, which produces differing forms of artefacts in WFS and Ambisonics [1], [15]. The frequency at which a linear or rectangular array of loudspeakers are half a wavelength apart is commonly termed the spatial Nyquist frequency of the array in the literature, in analogy with the sampling of continuous time signals [16]–[20]. Since surround sound typically entails sampling on a circle, for which equivalent theory exists, [21]–[23], the term is also applicable to surround systems.

The bandwidth limitations of surround sound systems based on the Ambisonics approach have been discussed in [14] and [24]. For example, reproduction of a 2D sound field over a

0.1 metre radius at frequencies up to 8 kHz requires the use of an array of at least 31 monopole sources. Reproduction of a 3D field over the same radius and bandwidth requires at least 250 monopole sources in a 3D array. For a circular array of 5 loudspeakers, as typically used in the home, the maximum frequency of holographic reproduction of 2D fields over a 0.1 metre radius is 1 kHz. At higher frequencies, various perceptual modifications are used to control artefacts that would be introduced by the application of the ideal low-frequency theory [25]–[29].

A second significant limitation is the effects of reflections of the loudspeaker sound from room surfaces and subsequent reverberation. While early reflections tend to produce an increase in apparent source width which produces a subjectively positive impression in concert halls [30][31], such effects are not desirable in a sound reproduction system whose goal is to faithfully recreate an existing ASW. Typically the short early reflections in the home will mask longer reflections from a larger space, giving a reduced impression of room size. Similarly, reverberation in a home situation tends to degrade the reproduction of the longer RT of a concert hall.

An alternative view of sound reproduction in the home is that the room provides additional ambience which can be subjectively beneficial. For example, omnidirectional speakers are preferred for stereo reproduction as they tend to provide early reflections which enhance stereo reproduction [32]. However, a surround system which is free from room reflections, and which has sufficient spatial resolution, can synthesise any desired early reflection sequence, giving greater control of subjective preference than can be provided by an individual room.

A potential solution to the two limitations described above is the use of higher order loudspeakers (HOLs), coupled with the application of active compensation to increase the reproduction bandwidth and radius and reduce the effects of room reflections [14], [33]–[38].

Higher order loudspeakers radiate sound with a number of different radiation patterns, which are typically described by spherical or cylindrical harmonics [39]–[43]. In the 2D case, polar responses in the horizontal plane are produced with variations in azimuthal angle ϕ of the form $\cos(n\phi)$ and $\sin(n\phi)$ [44][45]. For spherical loudspeakers the ideal response of these sources are the sectorial spherical harmonics [12][45]. For cylindrical loudspeakers, the response in elevation can be controlled by the vertical directivity of the radiating elements. For example, floor to ceiling loudspeakers can provide narrow beams in the (x,y) plane [46].

If active compensation is implemented, then a surround sound system can reduce the early reflections and some reverberation [33]–[37], [47]. The reduction must, however, occur over a region that exceeds the human head (a radius exceeding 8.75 cm [48]). It is commonly stated regarding equalization in reverberant environments that acoustic control is possible over one tenth of a wavelength [49]–[51], suggesting that reverberation-free reproduction is not possible at high frequencies. However, the one tenth wavelength result applies for the case of pressure control only. When higher order modes of the sound field are controlled the control radius can be extended [52]–[57].

This paper presents experimental results from a prototype 2D sound reproduction system using higher-order loudspeakers and a higher order microphone for room compensation. The number of HO loudspeakers is limited to five, since this number is common in commercial surround systems. The loudspeaker designs are based on circular arrays of drivers on cylindrical baffles as this is less complex than 3D designs using spherical baffles. Similarly

the microphone is a 2D cylindrical design which requires the sampling of less audio channels than a spherical microphone design.

It will be shown that the prototype sound system is able to attenuate early reflections and reverberation over a region of approximately 200 mm diameter up to a frequency of 8 kHz.

1 THEORY: ACTIVE COMPENSATION OF 2D SOUND FIELDS

The reproduction of 2D sound fields using the Ambisonics approach can be carried out using either a spherical harmonic approach known as the sectorial approximation [58], which requires the use of a spherical microphone array, or a cylindrical harmonic approach, which can be implemented using a cylindrical microphone array. As a cylindrical array was developed here, the corresponding cylindrical coordinate theory for determining the 2D component of the sound field and reproducing a desired 2D field is first presented.

1.1 Measurement of 2D sound field coefficients using a cylindrical microphone array

The spatial sound field at positive radian frequency ω incident upon a source-free region of space may be described in cylindrical coordinates (R,z,ϕ) by the interior solution to the wave equation [12]

$$p(R,z,\phi,\omega) = \sum_{m=-\infty}^{\infty} e^{im\phi} \frac{1}{2\pi} \int_{-\infty}^{\infty} A_m(k_z,\omega) J_m(k_R R) e^{-ik_z z} dk_z \,, \tag{1}$$

where $J_m(.)$ is the mth cylindrical Bessel function, $A_m(k_z,\omega)$ is the mth sound field coefficient which is a function of the wave number k and its z component $k_z = \sqrt{k^2 - k_R^2}$, where k_R is the radial part. The definition in (1) means that positive values of k_z are associated with wave front propagation in the positive z direction.

For a finite wave number k and radius R, Eq. (1) can be truncated to a maximum order $M = \lceil ekR/2 \rceil$ [59]. Far-field sound sources which are in the (x,y) plane have no propagation component in z and produce $k_z = 0, k_R = k,$ and so $A_m(k_z,\omega) = A_m(\omega)\delta(k_z)$, where $\delta(.)$ is the Dirac delta function with area $2\pi,$ and hence the sound pressure simplifies to

$$\overline{p}(R,\phi,\omega) = \sum_{m=-M}^{M} e^{im\phi} A_m(\omega) J_m(kR), \tag{2}$$

as is commonly assumed in 2D sound reproduction [10][25].

If the incident field in (1) scatters from an infinitely tall cylinder, the scattered field has the form [12]

$$p(R,z,\phi,\omega) = \sum_{m=-\infty}^{\infty} e^{im\phi} \frac{1}{2\pi} \int_{-\infty}^{\infty} B_m(k_z,\omega) H_m(k_R R) e^{-ik_z z} dk_z , \tag{3}$$

where $H_m(k_R R) = H_m^{(2)}(k_R R)$ is the cylindrical Hankel function of the second kind [12]. The total field is the sum of the incident and scattered fields (1) and (3). If the cylinder has infinite surface impedance then the normal pressure gradient on the surface must vanish. The scattering functions $B_m(k_z,\omega)$ can then be expressed in terms of $A_m(k_z,\omega)$, yielding the total field

$$p_T(R,z,\phi,\omega) = \sum_{m=-\infty}^{\infty} e^{im\phi} \frac{1}{2\pi} \int_{-\infty}^{\infty} A_m(k_z,\omega) \left[J_m(k_R R) - \frac{J_m'(k_R a)}{H_m'(k_R a)} H_m^{(2)}(k_R R) \right] e^{-ik_z z} dk_z . \tag{4}$$

On the surface, $R = a$, the application of the cylindrical Wronskian, [12], yields the pressure on the cylinder

$$p_T\left(a,z,\phi,\omega\right)=\frac{-2i}{\pi a}\sum_{m=-\infty}^{\infty}e^{im\phi}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{A_m\left(k_z,\omega\right)}{k_R H'_m\left(k_R a\right)}e^{-ik_z z}dk_z\ . \tag{5}$$

For a 2D sound reproduction system with loudspeakers in the horizontal plane, the direct field of each speaker, and the subsequent wall reflections, have $k_z=0$. If the ceiling and floor are absorptive, then reflections from the floor and ceiling are reduced and the reproduced field is predominantly 2D. In this case 2D sound field reproduction can be achieved by recording the 2D sound field component, with $k_z=0$, and reproducing a desired 2D field of the form of (2). The 2D field can be found by carrying out the integral in (5) with $k_z=0$, ie by integrating over z. In practice, integration is only possible over a finite range, and would typically be carried out by a discrete summation obtained from a line array of microphones. These practicalities can be taken into account by integrating with a general aperture function $f\left(z\right)$

$$p_T\left(a,\phi,\omega\right)=\int_{-\infty}^{\infty}p_T\left(a,z,\phi,\omega\right)f\left(z\right)dz\ , \tag{6}$$

yielding, from (5),

$$p_T\left(a,\phi,\omega\right)=\frac{-2i}{\pi a}\sum_{m=-\infty}^{\infty}e^{im\phi}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{A_m\left(k_z,\omega\right)}{k_R H'_m\left(k_R a\right)}F\left(k_z\right)dk_z\ , \tag{7}$$

where

$$F\left(k_z\right)=\int_{-\infty}^{\infty}f\left(z\right)e^{-ik_z z}dz \tag{8}$$

is the Fourier transform of the array aperture function $f\left(z\right)$. In practice, $f\left(z\right)$ is a discrete weighting function corresponding to a line array of N_L microphones in z, which produces an array response with a main lobe in the horizontal plane, $k_z=0$.

For sound sources that are at reasonable distances from the cylinder, the contribution of evanescent terms to the pressure are negligible and it can be assumed that the z-component of the wave number has magnitude less than k, and can be written $k_z = k\cos\theta$ with θ the elevation angle from the z-axis. In this case, Eq. (7) becomes

$$p_T\left(a,\phi,\omega\right) = \frac{-2i}{\pi a}\sum_{m=-\infty}^{\infty} e^{im\phi}\,\frac{1}{2\pi}\int_0^{\pi}\frac{A_m\left(k\cos\theta,\omega\right)}{H_m'\left(ka\sin\theta\right)}F\left(k\cos\theta\right)d\theta \tag{9}$$

$F\left(k\cos\theta\right)$ may now be interpreted as the polar response of the vertical aperture function at wave number k. With an absorptive floor and ceiling, and if $F\left(k\cos\theta\right)$ is sufficiently sharp to attenuate the effects of any residual vertical propagation, then the 2D coefficients, $A_m\left(0,\omega\right) = A_m\left(\omega\right)$, are obtained as

$$A_m\left(\omega\right) = \frac{i\pi a}{2}H_m'\left(ka\right)\frac{1}{2\pi}\int_0^{2\pi} p_T\left(a,\phi,\omega\right)e^{-im\phi}d\phi\,. \tag{10}$$

In practice, the integration over azimuthal angle ϕ is carried out using a discrete sampling at N_ϕ angles. The cylindrical microphone array thus consists of N_ϕ line arrays equally spaced around the cylinder, with each line array having N_L elements. The summation over z can be implemented in the analogue domain and the line array outputs sampled, so that the number of audio channels that must be sampled is only N_ϕ.

1.2 Generation of an actively compensated sound field

Equation (10) can be applied to the active calibration of a higher order sound reproduction system as follows. The sound system consists of a circular array of L higher order loudspeakers. Each loudspeaker is able to radiate $2N+1$ fields with polar responses in the azimuthal plane of the form $\cos\left(n\phi\right)$ and $\sin\left(n\phi\right)$. There are thus a total of $Q = L\left(2N+1\right)$

ways of exciting the sound field in the room. For the *qth* sound field excitation the impulse responses at the output of each line array, $g\left(q,\phi_v,t_l\right), v\in\left[1:N_\phi\right]$ is measured, where $t_l = l/f_s$ is the *l*th sample in time and f_s is the sample rate. The discrete Fourier transform of each impulse response is taken to produce the transfer function $G\left(q,\phi_v,\omega_l\right)$ and the discrete form of the circular decomposition in Eq. (10) carried out to produce the 2D coefficient functions

$$A_m\left(q,\omega_l\right) = \frac{i\pi a}{2} H'_m\left(k_l a\right) \frac{1}{N_\phi} \sum_{v=1}^{N_\phi} G\left(q,\phi_v,\omega_l\right) e^{-im\phi_v} \tag{11}$$

at the discrete set of frequencies ω_l.

The sound field to be reproduced at each ω_l has the form of Eq. (2), with coefficients $D_m\left(\omega_l\right)$

$$p_d\left(R,\phi,\omega_l\right) = \sum_{m=-M}^{M} e^{im\phi} D_m\left(\omega_l\right) J_m\left(k_l R\right). \tag{12}$$

At each frequency, the set of sound field coefficients for the *qth* room excitation is $A_m\left(q,\omega_l\right), q\in\left[1,Q\right]$. The sound field obtained from a weighted sum of these excitations, with weights w_q, has the form

$$\hat{p}\left(R,\phi,\omega_l\right) = \sum_{m=-M}^{M} e^{im\phi} J_m\left(k_l R\right) \sum_{q=1}^{Q} w_q\left(\omega_l\right) A_m\left(q,\omega_l\right). \tag{13}$$

Requiring this field to approximate (12) produces the mode matching requirements

$$\sum_{q=1}^{Q} w_q\left(\omega_l\right) A_m\left(q,\omega_l\right) = D_m\left(\omega_l\right). \tag{14}$$

For each frequency, this can be put in matrix form

$$\mathbf{\Psi} w = D\,, \tag{15}$$

where $\mathbf{\Psi}$ is of size $2M$+1 by Q . We assume that $Q \geq 2M+1$. In this case the weights may be found as

$$w = \mathbf{\Psi}^H \left[\mathbf{\Psi}\mathbf{\Psi}^H\right]^{-1} D\,. \tag{16}$$

The solution may produce weights with large energy which produces non-robust results. The energy of the weights can be reduced below that of the ideal minimum energy solution by including a regularization term λ

$$w_\lambda = \mathbf{\Psi}^H \left[\mathbf{\Psi}\mathbf{\Psi}^H + \lambda\mathbf{I}\right]^{-1} D \tag{17}$$

This equation is solved at each discrete frequency of interest. The frequency dependent, Q by $2M$+1 matrix $\mathbf{\Psi}^H \left[\mathbf{\Psi}\mathbf{\Psi}^H + \lambda\mathbf{I}\right]^{-1}$ produces a set of Q loudspeaker weights from a desired set of $2M$+1 sound field coefficients D and can be implemented using fast convolution techniques.

The maximum frequency at which the mode matching equations (14) can be achieved defines the spatial Nyquist frequency of the sound reproduction system for a given desired radius of reproduction R. Setting $Q = L\left(2N+1\right) = 2M+1$ with $M = \lceil ekR/2 \rceil$ yields

$$f_{Nyq}\left(N\right) = \frac{c\left[L\left(N+1\right)-1\right]}{2e\pi R} \tag{18}$$

For example, for a radius R = 90mm and N = 0 the spatial Nyquist frequency is 885 Hz and for N = 3 it is 5.6 kHz.

1.3 Active compensation over a finite duration

The active compensation of the entire room impulse response is a challenging task, requiring long filters and leading to possible ill-conditioning and non-robust reproduction. The problem complexity can be reduced by applying a window function to the measured impulse responses $g\left(q,\phi_v,t_l\right)$ which zeros the response for all samples after a truncation time τ after the direct sound, and determining the filter weights from the transfer functions of the windowed responses. In this case, active compensation produces a direct sound field with attenuated early reflections and early reverberation up to time τ, but there will be a subsequent reverberant field which adds some reverberation to the reproduced field.

2 DESCRIPTION OF PROTOTYPE 2D SURROUND SYSTEM USING HIGHER-ORDER LOUDSPEAKERS AND ACTIVE COMPENSATION

2.1 Higher order loudspeakers

The prototype loudspeaker is a third order cylindrical design consisting of 30 full range drivers arranged as two circular arrays on a 15-sided polygonal baffle with mean radius a = 132 mm, and height 400mm, plus a subwoofer mounted in the top which extends the omnidirectional mode response at low frequencies [45] (Fig 1).

Figure 1: Third order loudspeaker unit

The spatial Nyquist frequency of the loudspeaker is approximately 4 kHz. The use of two drivers at each angle provides increased directivity at high frequencies which reduces out-of-plane radiation lobes. The loudspeaker provides an omnidirectional response and higher-order 2D responses up to order 3, a total of 7 radiation modes. The *nth* radiation mode becomes significant for frequencies where $ka > n$, which defines the loudspeaker mode activation frequency

$$f_{\text{act,spkr}}(n) = \frac{nc}{2\pi a}, \tag{19}$$

which produces first to third order activation frequencies of 410, 820 and 1.2kHz, respectively. Each radiation mode is unable to radiate significant power below the corresponding activation frequency [41], [43], [60], [61]. This behaviour is well-known for the $n = 1$ dipole case [62].

2.2 Higher order microphone

Higher order microphones can be constructed using multiple microphones in a free-field geometry [63]–[67] or mounted on spherical or truncated cylindrical baffles [68]–[77]. The main design goals are to have a sufficient number of microphone elements to record the sound field up to a desired frequency and to record the sound field over a radius larger than the human head radius to allow an accurate perceived reproduction. The required radius of the microphone can be determined from Eq. (2). If the sound field is to be recorded and reproduced within a sound control zone (SCZ) of radius R and up to wavenumber k, then the expansion order of the sound field can be truncated to order $M \approx kR$ [78]. If the recording is carried out using a cylindrical or spherical microphone array of radius b, then the microphone array must have significant output at mode order M. Since the activation frequency for the mth mode of a cylindrical or spherical array of radius b occurs for $kb = m$ [77],

$$f_{\text{act,mic}}(m) = \frac{mc}{2\pi b}, \tag{20}$$

then for a sound field order $m = M$, its radius must be at least $b = R$. For sound field reproduction over a radius exceeding the human head, a design radius exceeding 87.5 mm [48] is therefore required.

The prototype cylindrical microphone consists of a circular array of 32 line arrays positioned on a cylinder of radius 90 mm (Fig 2). Each line array consists of 5 MEMs microphone elements, spaced at -60, -20, 0, 20 and 60mm from the vertical center. The outputs of the 5 microphones are filtered and added together to provide an estimate of the 2D component of the sound field as discussed in section 1.1. The filtering is included to maintain a more constant directivity in elevation with frequency [79]. The total number of microphones is 165, but the line arrays are combined, and the cylindrical array has 32 analog outputs which

can be sampled using commercially available analog to digital converters. For comparison, the number of microphones and analogue to digital converter channels required for a spherical microphone array of the same radius and maximum recording frequency would be 205.

The cylindrical array is able to measure 2D cylindrical sound field coefficients up to 15th order and up to 8 kHz. In practice, the accuracy of the highest orders is limited and so the order was restricted to a maximum of 12. This allows the recording and reproduction of sound fields over a radius of 90 mm up to 7.2 kHz.

Figure 2: Higher order microphone

2.3 Reproduction room.

The reproduction space was a rectangular office of dimensions 5.10 m length by 4.55 m width by 3.22 m height, with a carpeted floor. The right hand wall, with respect to the 0 degrees direction, was largely covered by windows of dimensions 4.8 m by 2.05 m and the opposite side had a small window and a power box, creating some diffusion in the lateral dimension. The ceiling was treated with 50mm acoustic insulation with a 50mm air gap to reduce vertical reflections [80]. The reverberation time with acoustic treatment was 450 ms. Some measurements were taken before the ceiling absorption was installed, and these are discussed in section 3.

The layout of the sound system is shown in Fig. (3). The HOLs are arranged in a regular pentagonal array at a radius of 1.5 m. The regular layout allowed comparison of the results with standard Ambisonics approaches. Both the HOLs and the calibrating microphones were located at 1.1 m above the ground.

Figure 3: Reproduction room with surround sound layout

2.4 Audio processing

Audio sampling and playback was carried out using an RME data acquisition system with 32 channels of audio input and 64 channels of audio output. The 32 input channels were used to sample the cylindrical microphone outputs and 35 of the 64 output channels used to feed signals to the 5 HOLs. The data acquisition system was connected to a computer via a MADI serial link. The surround calibration and playback was carried out in Matlab using the Playrec audio utility [81].

Calibration was carried out by measuring the impulse responses from each radiation mode of each HOL to the microphones in the microphone array. This was done by generating a periodically repeated chirp sequence whose periodic spectrum (calculated via an FFT) was spectrally flat. The chirp was designed in the frequency domain in a similar manner to those

developed in [82]. The chirp length was 44100 samples at a sample rate of 44.1 kHz giving a period of 1 second. For each channel of each HOL, the repeated chirp was played and the 32 microphone array output signals were recorded. After the first period, or 'frame', of the chirp, the response becomes periodic, and the second and subsequent periods are used to calculate the transfer function by taking the FFT of one frame and dividing it by the FFT of the chirp. Multiple frames may be added to improve the signal to noise ratio, assuming time invariance. In practice a single repeat was used which provided 3 dB improvement in SNR and typical overall SNRs were around 65 dB. The calibration of the entire system took three minutes.

Actively compensated sound field reproduction (AC-SFR) was carried out by calculating the 2D cylindrical harmonic decomposition of the microphone outputs using Eq. (11) at each frequency for each loudspeaker excitation, and determining the weights required to produce a desired 2D sound field, as in Eq. (12), using Eq. (17). In practice a weighted form of mode-matching was used to design the compensating loudspeaker filters [83]. A Tikhonov parameter of 1/100 of the largest singular value in $\boldsymbol{\Psi}$ was applied in (17) in calculating the loudspeaker filters. Due to the bandwidth limitations of the loudspeaker and microphones, reproduction was restricted to a maximum frequency of 8 kHz.

2.5 Analysis of reproduced field

The reproduced field was assessed by informal listening tests and by measurement of impulse responses using an Eigenmike™ , a 32-element spherical microphone array.

For informal listening, the source signal was a short burst of white noise or a section of anechoically recorded music. A range of truncation times from τ = 30ms to 500 ms were assessed.

For the objective analysis, the reproduced sound field was an impulse response and active compensation was applied over a time duration $\tau = 50$ ms after the direct sound. This truncation time is equivalent to including up to approximately third order reflections.

To improve the signal to noise ratio, the impulse responses were calculated by using the repeating chirp as the desired source signal for a desired reproduction angle. The reproduced field was recorded using the Eigenmike and processed as discussed in section 2.4 to produce the corresponding impulse responses at each microphone. The impulse response at the center of the Eigenmike was calculated by adding all 32 outputs and applying a zeroth order sphere mode equalization.

The spatial properties of the sound field at the Eigenmike position were determined by produce a maximum directivity beam for a discrete set of angles (θ_l, ϕ_l) over 4π steradians, determining the total energy for each and then producing a 3D Steered Response Power (SRP) which shows the energy arriving from each direction. The beam look for each direction has a normalized polar response of the form [61]

$$d_V(\theta,\phi) = \frac{4\pi}{(V+1)^2} \sum_{v=0}^{V} \sum_{m=-v}^{v} Y_v^m(\theta,\phi) Y_v^m(\theta_l,\phi_l)^* \tag{21}$$

where V is the beam order, which is dictated by the activation frequency for the Eigenmike, as in Eq. (20). With the use of the addition theorem, Eq. (21) can be written

$$d_V(\gamma) = \frac{1}{(V+1)^2} \sum_{v=0}^{V} (2v+1) P_v(\cos\gamma) \tag{22}$$

where γ is the angle between (θ_l, ϕ_l) and (θ, ϕ). This response can be produced using the Eigenmike™ for frequencies above the *Vth* activation frequency up to a maximum order $V = 4$. The beamformer steered response power was measured in frequency bands centred

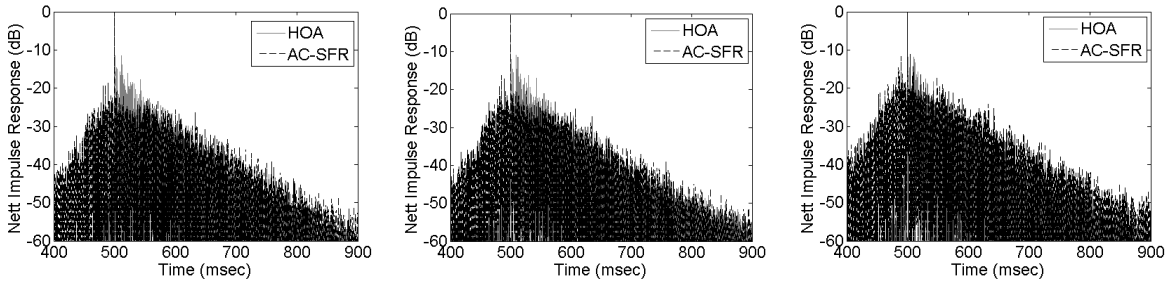
about 250 Hz, 500 Hz, 1 kHz, 2 kHz and 4 kHz. Due to the increase in directivity of beamformers with frequency, these SRPs naturally look sharper in higher frequency bands.

3 RESULTS

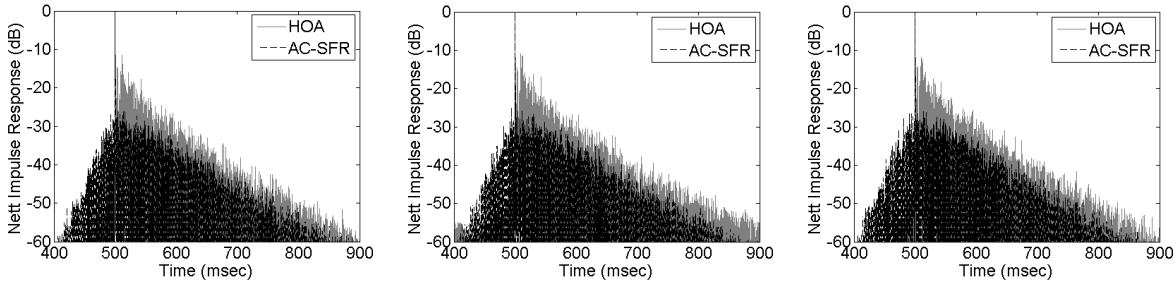
The surround system was evaluated by measuring actively-compensated impulse responses at the centre of the array and at 45mm and 90mm lateral displacement from the center. The spatial properties of the field were evaluated at the array center.

Actively-compensated impulse responses were also reproduced using only the zeroth order modes of the five loudspeakers, in order to determine if there is any performance improvement due to the use of higher order loudspeakers. In addition, impulse responses were recorded using HOA with frequency-independent second order max-r_E decoding [25]. Unless otherwise stated, measurements were taken in the room with ceiling absorption.

Figures 4 and 5 show the reproduced impulse responses for the creation of sources at 0 and 36 degrees (half way between two loudspeakers), respectively, for AC-SFR using zeroth order (omnidirectional) and third order loudspeakers, together with the results for HOA. All reproduced impulse responses were time aligned and normalised to direct sound levels of 0 dB for comparison.

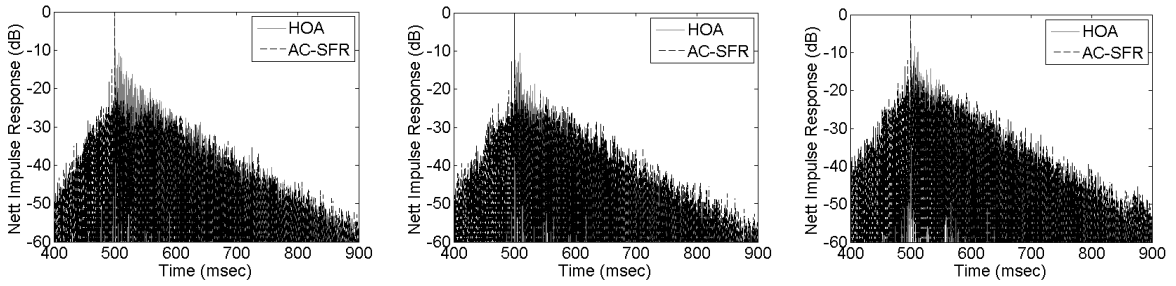


(a) centre, 0th order **(b) 45 mm, 0th order** **(c) 90 mm, 0th order**

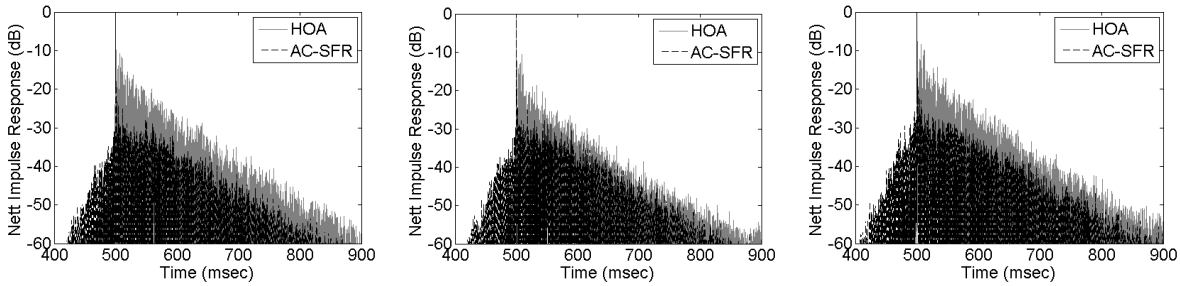


(d) centre, 3rd order **(e) 45 mm, 3rd order** **(f) 90 mm, 3rd order**

Figure 4: Nett impulse responses at several positions in the reproduction zone, comparing AC-SFR using (a),(b),(c) zeroth order secondary sources and (d),(e),(f) third order secondary sources with second order Ambisonics (max-r_E) for a virtual source angle of 0 degrees.



(a) centre, 0th order **(b) 45 mm, 0th order** **(c) 90 mm, 0th order**



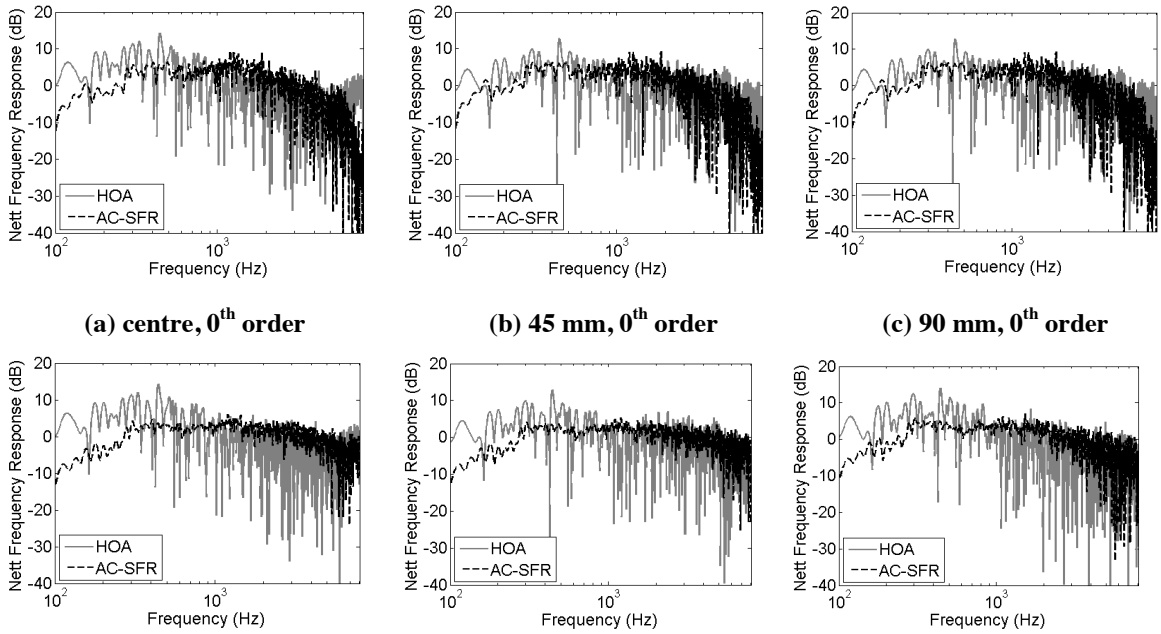
(d) centre, 3rd order **(e) 45 mm, 3rd order** **(f) 90 mm, 3rd order**

Figure 5: Nett impulse responses at several positions in the reproduction zone, comparing AC-SFR using (a),(b),(c) zeroth order secondary sources and (d),(e),(f) third order secondary sources with second order Ambisonics (max-r_E) for a virtual source angle of 36 degrees.

The third-order AC-SFR approach is seen to suppress most of the early reflections within the first 50 ms, with an attenuation of about 30 dB for both 0 and 36 degree source angles, and at all three positions in the SCZ. Using AC-SFR with zeroth order sources, most of the early reflections are suppressed by about 20 dB. For the HOA case the reflections are about 10 dB below the direct component. Hence, zeroth order AC-SFR produces about 10 dB reduction of reflections and third-order AC-SFR produces 20 dB attenuation, compared to the un-compensated HOA case.

The undesirable feature of the reproduced impulse responses with AC-SFR is the presence of "pre-reverberation", where there is a build-up of sound energy before the desired impulse. The pre-reverberation in Figure 5 rises to a maximum level of –20 dB before the direct sound for zeroth order sources and to –30 dB for third order sources. It is hence about 10 dB lower with third order loudspeakers than it is for zeroth order sources.

If the pre-reverberation duration is short enough and at sufficiently low level, it will not be audible due to the perceptual pre-masking effect [84]. However, the pre-reverberation with third-order sources still exceeds the threshold for pre-masking [84], and was clearly audible when listening to white noise pulses in informal listening tests.



(a) centre, 0th order **(b) 45 mm, 0th order** **(c) 90 mm, 0th order**

(d) centre, 3rd order **(e) 45 mm, 3rd order** **(f) 90 mm, 3rd order**

Figure 6: Nett frequency responses at several positions in the reproduction zone, comparing AC-SFR using (a),(b),(c) zeroth order secondary sources and (d),(e),(f) third order secondary sources with second order Ambisonics (max-r_E) for a virtual source angle of 36 degrees.

In Figure 6, the nett reproduced frequency responses at the three locations are shown for the different reproduction approaches[1]. AC-SFR produces flattened frequency responses, corresponding to the attenuation of early reflections and early reverberation, with the frequency responses being most flat in the midrange 300 Hz – 1 kHz. Zeroth order reproduction suppresses the appearance of significant nulls in the frequency responses until about 1 kHz and third order reproduction extends the region of flatness to around 3 kHz. The frequency responses for zeroth order reproduction roll off above 2 kHz. This occurs because the five channel system can only sustain reproduction to about 1.2 kHz and the reverberant compensation problem becomes ill-conditioned here so that the reproduced sound energy is reduced by regularization.

[1] To study the effectiveness of the AC-SFR approaches, all frequency responses are plotted for only the first 50 msec of reproduced impulse responses.

Table 1 quantifies the level of reverberation suppression more directly, by showing the direct to reverberant ratio (DRR) for the three approaches at the three positions, averaged over ten virtual source directions 0^o, 36^o, … 324^o. (For AC-SFR the DRR calculation included the pre-reverberation). Third order AC-SFR suppresses more reverberation than the zeroth order case. Both AC-SFR scenarios exceed the DRR performance of HOA, which can be considered the baseline. The DRRs of third and first order sources average to 6.4 dB and 1.5 dB respectively, whilst those of HOA average to −2.5dB. On average, the third-order DRR is about 5 dB higher than the zeroth-order DRR.

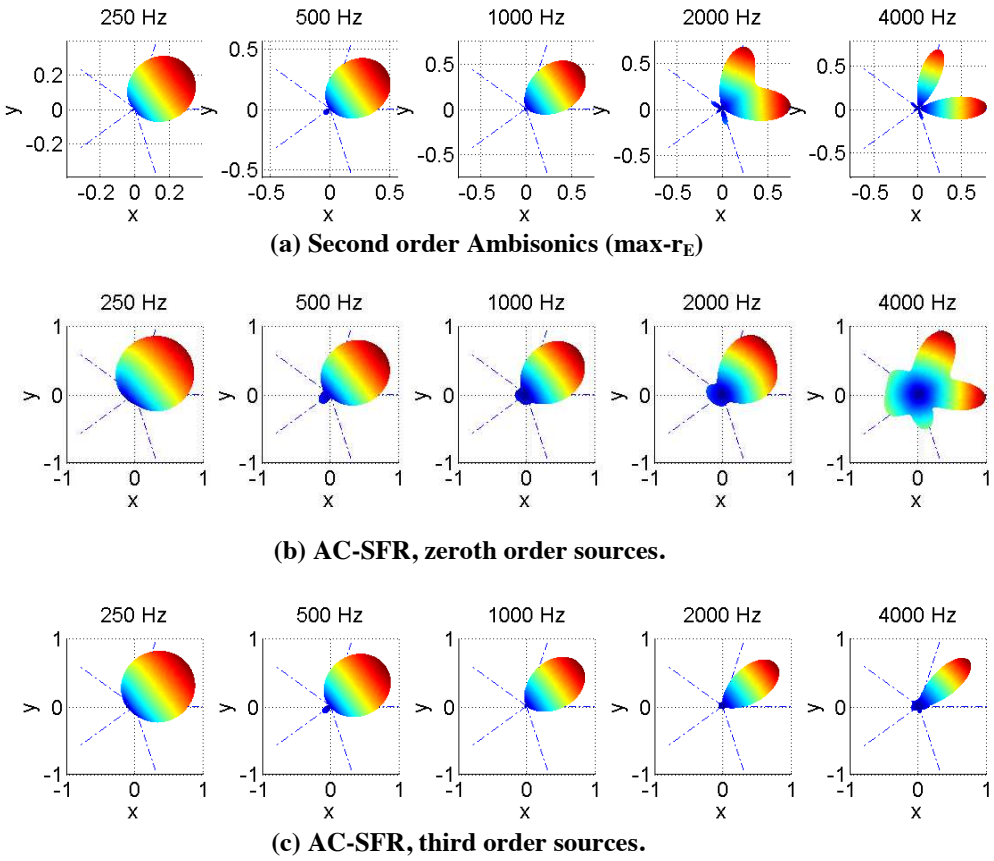
Position	Direct-to-Reverberant energy Ratio (dB)		
	AC-SFR, 3^{rd} order	AC-SFR, 0^{th} order	HOA (max-r_E)
Centre	6.7	2.5	-2.6
45 mm	6.4	2.1	-2.4
90 mm	5.4	0.03	-2.4

Table 1: Performance for second-order Ambisonics (max-r_E), AC-SFR using 0^{th} order modes and AC-SFR using 3^{rd} order modes, showing the Direct-to-Reverberant energy Ratio of the nett impulse responses averaged over ten virtual source angles.

Figure 7 shows the beamformer steered response power at the centre of the SCZ for a source angle of 36 degrees. At low frequencies, HOA produces maximum energy from halfway between the loudspeakers, but at 2 kHz and above produces two separate energy peaks at the two nearest loudspeaker angles, showing that the system is unable to produce a single direction of arrival at these frequencies. This is expected, because for *N=2* order HOA the size of the SCZ is smaller than the radius of the Eigenmike (42 mm) for frequencies above 4.5 kHz.

The zeroth order AC-SFC produces some improvement over the HOA case. The 2 kHz energy has a single direction of arrival although the mean is slightly broader than the 1 kHz case. However at 4 kHz, the energy arrives from two distinct directions.

The third-order AC-SFC produces a consistent beam up to 4 kHz and the beam width is narrowest at 4 kHz, showing that sound field reproduction is correct up to 4 kHz over the radius of the Eigenmike.



(a) Second order Ambisonics (max-r_E)

(b) AC-SFR, zeroth order sources.

(c) AC-SFR, third order sources.

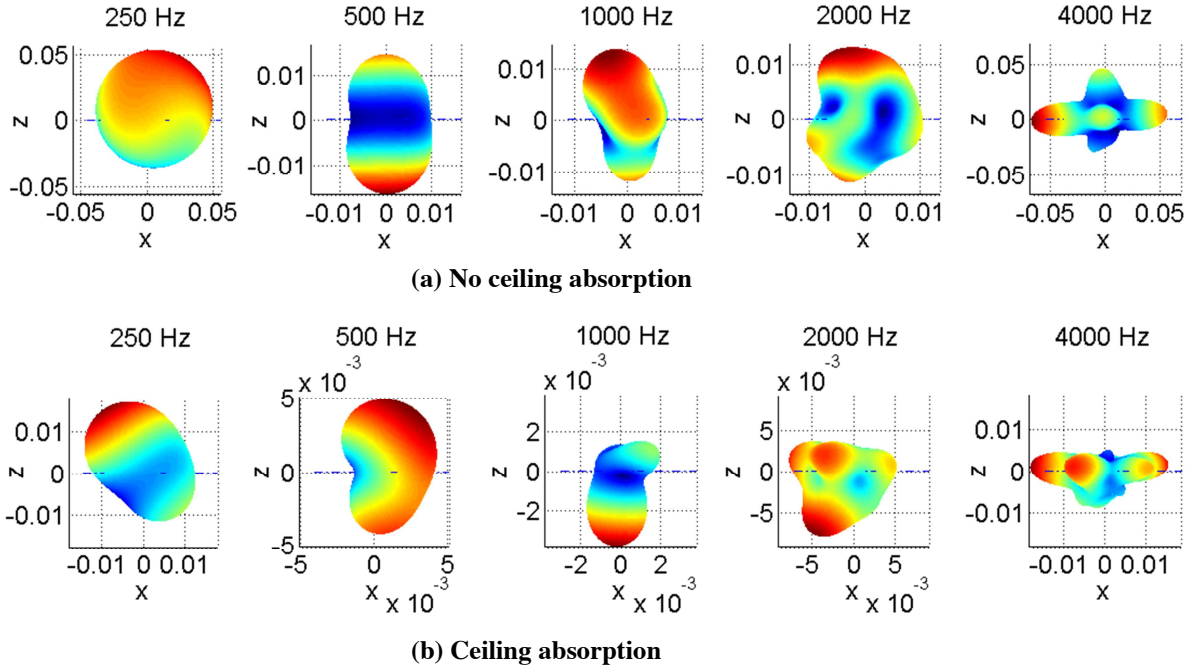
Figure 7: Beamformer SRP at the centre of the sound control zone in various frequency bands for (a) AC-SFR with zeroth order sources, (b) AC-SFR with third order sources and (c) Second order Ambisonics. Plots are shown for a virtual source angle of 36 degrees. The directions to the secondary sources are also marked (-·).

The spatial properties of the pre-reverberation signal can be determined by windowing the measured impulse response so that the desired direct sound and subsequent compensated response is removed, leaving only the pre-reverberation build-up. In practice the window was arranged to select the impulse response up to 2 ms before the direct sound.

The SRP of the pre-reverberation component is shown in Fig. 8a without ceiling absorption, viewed in the *x*-*z* plane. These measurements used a 100 ms truncation time as opposed to the

50 ms time used above, which creates slightly higher levels of pre-reverberation. Vertical lobes are evident at 500 Hz, 1 kHz and 2 kHz. At 4 kHz the higher order loudspeakers are more directional so that vertical lobes do not dominate the responses.

In Fig. 8b the SRP of the pre-reverberation is shown after the application of ceiling absorption. The upward vertical lobes at 1, 2 and 4 kHz are now largely absent.



(a) No ceiling absorption

(b) Ceiling absorption

Figure 8: Beamformer SRP of the pre-reverberation in several frequency bands for a virtual source at 36 degrees, midway between two loudspeakers, for a room with 50mm ceiling absorption.

The frequency dependence of the pre-reverberation levels is shown in Fig. 9. The pre-reverberation levels of 3rd order sources are around 10 dB lower than the levels for 0th order sources, as suggested by Figures 5 and 6.

If the pre-reverberation properties are the same as that of room reverberation, the levels would rise at 10 dB per decade on a log frequency scale, but Fig. 9 shows that they rise at greater than 10 dB per decade. This increasing pre-reverberation level with frequency is

likely to be a consequence of the poorer reproduction accuracy that occurs at high frequencies.

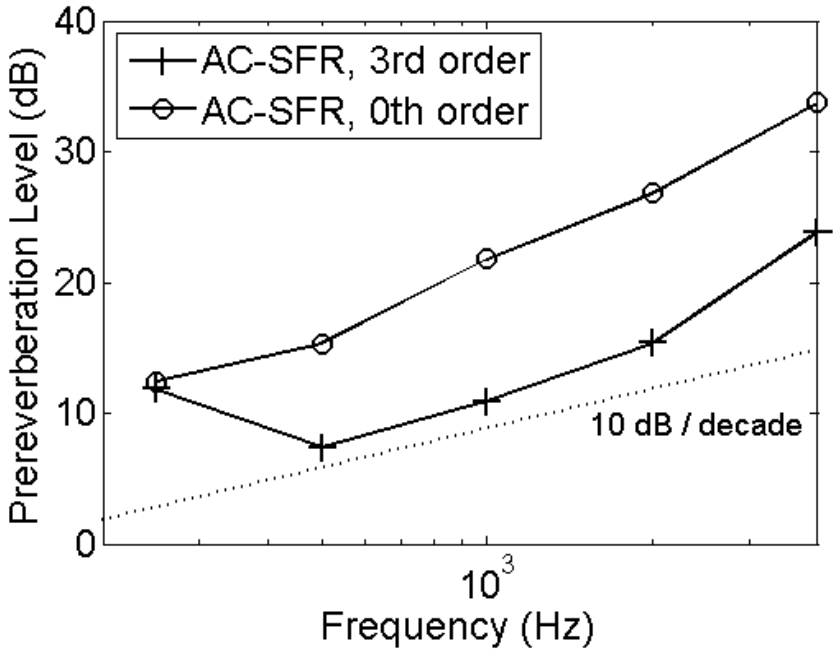


Figure 9: Relative pre-reverberation levels in octave bands versus frequency.

4 CONCLUSION

In this paper we have considered the generation of actively compensated 2D sound fields in a semi-reverberant room, using an array of five third-order loudspeakers which can exploit room reflections to improve reproduction quality. Both loudspeakers and the calibration microphone were cylindrical designs. Recording and playback was carried out using commercial data acquisition equipment and software.

The results have shown that active compensation together with the use of higher order loudspeakers can produce sound fields with attenuated early reflections and early reverberation over a radius of 90 mm, large enough for a single listener, and up to a frequency of 8 kHz. Furthermore, third order sources produced greater suppression of early reflections than zeroth order sources and the frequency response was equalised over a wider

frequency range. Eq. (18) predicts that the spatial Nyquist frequency of the zeroth order system is 885 Hz and is around 5.6 kHz for the third order case. A definitive analysis of the spatial Nyquist frequency would require sampling the sound field over the reproduction region, which has not been carried out here. However, the results in Figures 6 and 7 clearly show that the frequency range of accurate reproduction is extended by using higher order sources.

AC-SFR also produced an unwanted pre-reverberation response. This pre-reverberation is caused by any vertical propagating waves that the 2D system cannot control, and by any errors in the 2D mode matching caused by measurement errors and any variation of the transfer functions that occurs after the calibration process. For example, replacing the calibration microphone with a human listener will alter the secondary room reflections caused by the field that diffracts around the microphone.

The vertical propagation component of the pre-reverberation was reduced by using ceiling absorption to attenuate unwanted vertically propagating wave fronts.

The pre-reverberation levels were around 10 dB lower using third-order sources than using zeroth order sources and tended to increase with frequency. This suggests that using sources of order higher than three may further attenuate the pre-reverberation signal by providing additional room excitations to control higher-order modes at high frequencies. The activation frequency of the third order mode of the prototype speaker is 1.23 kHz, and no higher radiation modes are available to increase the number of room excitations above this frequency. Higher order radiation modes would provide greater capacity to reduce pre-reverberation up to the maximum reproduction bandwidth, which was 8 kHz in this case.

Pre-reverberation may also be reduced by the use of a spherical microphone array, which would avoid any errors introduced by the use of a finite length cylindrical baffle and any

limitations caused by the finite vertical line arrays to record the 2D components of the sound field. However a spherical array would require the use of the sectorial approximation for 2D reproduction and so there may still be limitations for 2D reproduction with this approach. A difficulty with the spherical array is that a much large number of microphones and analogue to digital convertors would be needed. For the reproduction of sound fields up to 8 kHz, a microphone spherical array of radius 90 mm would require the sampling of over 205 microphone outputs.

Finally, the use of 3D higher order sources which can direct sound in the vertical dimension, together with 3D cylindrical or spherical microphone arrays, would provide greater control of any residual vertical propagation. Spherical sources, or cylindrical sources with line arrays that can beamform to produce vertically propagating wave fronts, could be used for 3D reproduction. However, the cost and complexity of such sources would be higher than for the 2D case.

The results in this paper did not take into account temperature variations in the reproduction room. Measurements were taken immediately after calibration to avoid the effects of any temperature changes. Temperature changes lead to a change in the speed of sound which alters the number of wavelengths along the propagation path. Since the acoustic wavelength is shorter at higher frequencies, temperature susceptibility is also more apparent at higher frequencies, and may contribute to the increase in pre-reverberation seen at high frequencies. For early reflections, the sound travels a shorter distance to reach the sound control region and the phase shift associated with a change in speed of sound is also smaller. Therefore truncation of the room impulse responses before active compensation tends to improve the robustness of the approach to temperature changes and was used in this paper. If compensation is attempted over a longer length of the room impulse responses, then

temperature compensation is necessary [85]. Alternatively, a microphone array must be in-place all the time and an adaptive system must be implemented such as in [35] [86][87].

The analysis of the performance of the system has so far been limited to objective measurements of the spatial properties of the reproduced sound fields using a spherical microphone array and informal listening. More detailed measurements, such as discussed in [88], have not been carried out. Such tests would be important to establish the perceived benefit of the ability to cancel reverberation and to create sound fields with arbitrary spatial properties, particularly since continuous sound with multiple directions will tend to average out and mask the effects of any residual pre-reverberant components in the reproduced field.

The informal subjective impression of the reproduced field for anechoic white noise pulses is that the direct sound appears to originate from a point immediately adjacent to the head. This impression is consistent with that produced in an anechoic environment. In contrast, standard second-order Ambisonics playback produces sound that appears to originate at a distance from the head, due to the presence of early reflections. Hence, the compensated surround system offers to prospect of the ability to alter the perceived distance of the sound source.

The pre-reverberation signal was also audible with white noise pulses and contributed to the residual reverberant field when listening to anechoic music. Methods of reducing the pre-reverberation field are discussed above and will be the focus of future research.

5 REFERENCES

[1] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter, "Spatial Sound With Loudspeakers and Its Perception: A Review of the Current State," *Proc. IEEE*, vol. 101, no. 9, pp. 1920–1938, 2013.

[2] J. Merimaa and V. Pulkki, "Spatial Impulse Response Rendering I: Analysis and Synthesis," *J. Audio Eng. Soc.*, vol. 53, no. 12, pp. 1115–1127, 2005.

[3] F. Baumgarte and C. Faller, “Binaural cue coding - Part 1: Psychoacoustic fundamentals and design principles,” *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, pp. 509–519, 2003.

[4] C. Faller and F. Baumgarte, “Binaural cue coding - Part II: Schemes and applications,” *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, pp. 520–531, 2003.

[5] A. J. Berkhout, D. de Vries, and P. Vogel, “Acoustic control by wave field synthesis,” *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, 1993.

[6] M. M. Boone, E. N. G. Verheijen, and P. F. Van Tol, “Spatial sound-field reproduction by wave-field synthesis,” *J. Audio Eng. Soc.*, vol. 43, no. 12, pp. 1003–1012, 1995.

[7] S. Spors, R. Rabenstein, and J. Ahrens, “The Theory of Wave Field Synthesis Revisited,” in *AES 124th Convention*, 2008, vol. 124.

[8] M. A. Gerzon, “Periphony: With-height sound reproduction,” *J. Audio Eng. Soc.*, vol. 21, no. 1, pp. 2–10, 1973.

[9] M. A. Gerzon, “Ambisonics in multichannel broadcasting and video,” *J. Audio Eng. Soc.*, vol. 33, no. 11, pp. 859–871, 1985.

[10] J. S. Bamford and J. Vanderkooy, “Ambisonics sound for us,” in *AES 99th Convention*, 1995.

[11] J. Daniel, “Representation de champs acoustiques, application a la transmission et a la reproduction de scenes sonores complexes dans un contexte multimedia,” University of Paris, 2000.

[12] E. G. Williams, *Fourier Acoustics*. San Deigo: Academic Press, 1999, pp. 115–234.

[13] R. Nicol and M. Emerit, “3D-sound reproduction over an extensive listening area: A hybrid method derived from Holophony and Ambisonic,” in *AES 16th International Conference*, 1999.

[14] M. A. Poletti, T. Betlehem, and T. D. Abhayapala, “Higher Order Loudspeakers for Improved Surround Sound Reproduction in Rooms,” in *AES 133rd convention*, 2012.

[15] J. Ahrens, H. Wierstorf, and S. Spors, “Comparison of higher order Ambisonics and wave field synthesis with respect to spatial discretization artifacts in time domain,” in *AES 40th Int. Conf*, 2010.

[16] P. de Groot and L. Deck, “Three-dimensional imaging by sub-Nyquist sampling of white-light interferograms,” *Opt. Lett.*, vol. 18, no. 17, pp. 1462–1464, 1993.

[17] M. D. Zoltowski and C. P. Matthews, “Real-time frequency and 2-D angle estimation with sub-Nyquist spatio-temporal sampling,” *IEEE Trans. Signal Proc.*, vol. 42, no. 10, pp. 2781–2794, 1994.

[18] B. G. Ferguson, "Remedying the Effects of Array Shape Distortion on the Spatial Filtering of Acoustic Data from a Line Array of Hydrophones," *IEEE J. Ocean. Eng.*, vol. 18, no. 4, pp. 565–571, 1993.

[19] J. Dmochowski, J. Benesty, and S. Affes, "On spatial aliasing in microphone arrays," *IEEE Trans. Signal Proc.*, vol. 57, no. 4, pp. 1383–1395, 2009.

[20] T. J. Rouphael and J. R. Cruz, "A spatial interpolation algorithm for the upsampling of uniform linear arrays," *IEEE Trans. Signal Proc.*, vol. 47, no. 6, pp. 1765–1769, 1999.

[21] H. Stark and M. Wengrovitz, "Comments and corrections on the use of polar sampling theorems in CT," *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. 31, no. 5, pp. 1329–1331, 1983.

[22] E. Margolis and Y. C. Eldar, "Reconstruction of nonuniformly sampled periodic signals: algorithms and stability analysis," in *Proc. 11th IEEE Int. Conf. Electronics, Circuits and Systems*, 2004, pp. 555 – 558.

[23] M. A. Poletti, "The design of encoding functions for stereophonic and polyphonic sound systems," *J. Audio Eng. Soc.*, vol. 44, no. 11, pp. 948–963, 1996.

[24] S. Bertet, J. Daniel, E. Parizet, and O. Warusfel, "Investigation on localisation accuracy for first and higher order Ambisonics reproduced sound sources," *Acta Acust. united with Acust.*, vol. 99, pp. 642–657, 2013.

[25] J. Daniel, J.-B. Rault, and J.-D. Polack, "Ambisonics encoding of other audio formats for multiple listening conditions," in *AES 105th Convention*, 1998.

[26] P. G. Craven, "Continuous surround panning for 5-speaker reproduction," in *AES 24th International Conference*, 2003.

[27] B. Wiggins, "The generation of panning laws for irregular speaker arrays using heuristic methods," in *AES 31st International Conference*, 2007.

[28] D. Moore and J. Wakefield, "The Design of Ambisonic Decoders for the ITU 5.1 Layout with Even Performance Characteristics," in *AES 124th Convention*, 2008.

[29] D. M. Murillo, F. M. Fazi, and M. Shin, "Evaluation of ambisonics decoding methods with experimental measurements," in *Proc. EAA Joint Symposium on Auralization and Ambisonics*, 2014.

[30] L. L. Beranek, *Concert and Opera Hall, How They Sound*. Published for the Acoustical Society of America by the American Institute of Physics, 1996.

[31] A. H. Marshall and M. Barron, "Spatial responsiveness in concert halls and the origins of spatial impression," *Appl. Acoust.*, vol. 62, pp. 91–108, 2001.

[32] W. Evans, J. Dyreby, S. Bech, S. Zielinski, and F. Rumsey, "Effects of loudspeaker directivity on perceived sound quality - a review of existing studies," in *AES 126th Conv.*, 2009.

[33] T. Betlehem and M. A. Poletti, "Two dimensional sound field reproduction using higher order sources to exploit room reflections," *J. Acoust. Soc. Am.*, vol. 135, no. 4, pp. 1820–1833, 2014.

[34] P.-A. Gauthier and A. Berry, "Adaptive Wave Field Synthesis for Sound Field Reproduction: Theory, Experiments, and Future Perspectives," *J. Audio Eng. Soc.*, pp. 1–22, 2007.

[35] S. Spors, H. Buchner, R. Rabenstein, and W. Herbordt, "Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering.," *J. Acoust. Soc. Am.*, vol. 122, no. 1, pp. 354–369, 2007.

[36] M. Kolundzija, C. Faller, and M. Vetterli, "Reproducing Sound Fields Using MIMO Acoustic Channel Inversion," *J. Audio Eng. Soc.*, vol. 59, no. 10, pp. 721–734, 2011.

[37] E. Corteel, "Equalization in an Extended Area Using Multichannel Inversion and Wave Field Synthesis," *J. Audio Eng. Soc.*, vol. 54, no. 12, pp. 1140–1161, 2006.

[38] A. Canclini, L. Markovic, F. Bianchi, F. Antonacci, A. Sarti, and S. Tubaro, "A geometric approach to room compensation for sound field rendering applications," in *Proc. 22nd European Signal Proc. Conf.*, 2014.

[39] R. Avizienis, A. Freed, and P. Kassakian, "A compact 120 independent element spherical loudspeaker array with programmable radiation patterns," in *AES 120th Convention*, 2006.

[40] O. Warusfel, P. Derogis, and R. Causse, "Radiation Synthesis with Digitally Controlled Loudspeakers," in *AES 103rd Convention*, 1997.

[41] A. M. Pasqual, J. R. De França Arruda, and P. Herzog, "Application of Acoustic Radiation Modes in the Directivity Control by a Spherical Loudspeaker Array," *Acta Acust. United With Acust.*, vol. 96, no. 1, pp. 32–42, 2010.

[42] M. Pollow and G. K. Behler, "Variable Directivity for Platonic Sound Sources Based on Spherical Harmonics Optimization," *Acta Acust. United With Acust.*, vol. 95, no. 6, pp. 1082–1092, 2009.

[43] B. Rafaely, "Spherical loudspeaker array for local active control of sound.," *J. Acoust. Soc. Am.*, vol. 125, no. 5, pp. 3006–3017, 2009.

[44] M. Kolundzija, C. Faller, and M. Vetterli, "Design of a compact cylindrical loudspeaker array for spatial sound reproduction," in *AES 130th convention*, 2011.

[45] M. A. Poletti and T. Betlehem, "Design of a prototype variable directivity loudspeaker for improved surround sound reproduction in rooms," in *AES 52nd Intl. Conf.*, 2013.

[46] R. D. White, "Wide-Range Electrostatic Loudspeaker with a Zero-Free Polar Response," *J. Audio Eng. Soc.*, vol. 57, no. 10, pp. 822–831, 2009.

[47] L.-J. Brännmark, A. Bahne, and A. Ahlén, "Compensation of Loudspeaker–Room Responses in a Robust MIMO Control Framework," *IEEE Trans. Audio, Speech Lang. Proc.*, vol. 21, no. 6, pp. 1201–1216, 2013.

[48] G. F. Kuhn, "Model for the interaural time differences in the azimuthal plane," *J. Acoust. Soc. Am.*, vol. 62, no. 1, pp. 157–167, 1977.

[49] S. J. Elliott, P. Joseph, A. J. Bullmore, and P. A. Nelson, "Active cancellation at a point in a pure tone diffuse sound field," *J. Sound Vib.*, vol. 120, no. 1, pp. 183–189, 1988.

[50] B. Rafaely, "Zones of quiet in a broadband diffuse sound field," *J. Acoust. Soc. Am.*, vol. 110, no. 1, pp. 296–302, 2001.

[51] B. D. Radlovic, R. C. Williamson, and R. A. Kennedy, "Equalization in an acoustic reverberant environment: robustness results," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 3, pp. 311–319, 2000.

[52] J. Daniel, "Spatial sound encoding including near field effect: Introducing distance coding filters and a viable new ambisonics format," in *AES 23rd International Conference*, 2003.

[53] M. A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Soc.*, vol. 53, no. 11, pp. 1004–1025, 2005.

[54] D. S. Talagala, W. Zhang, and T. D. Abhayapala, "Efficient multi-channel adaptive room compensation for spatial soundfield reproduction using a modal decomposition," *J. Acoust. Soc. Am.*, vol. 22, no. 10, pp. 1522–1532, 2014.

[55] B. Xu and S. Sommerfeldt, "Generalized acoustic energy density based active noise control in single frequency diffuse sound fields," *J. Acoust. Soc. Am.*, vol. 136, no. 3, pp. 1112–1119, 2014.

[56] A. Barkefors, S. Berthilsson, and M. Sternad, "Extending the area silenced by active noise control using multiple loudspeakers," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, 2012, pp. 325–328.

[57] A. Barkefors, M. Sternad, and L.-J. Brannmark, "Design and analysis of linear quadratic Gaussian feedforward controllers for active noise control," *IEEE/ACM Trans. Audio, Speech, Lang. Proc.*, vol. 22, no. 12, pp. 1777–1791, 2014.

[58] J. Ahrens and S. Spors, "An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions," *Acta Acust. united with Acust.*, vol. 94, pp. 988–999, 2008.

[59] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Trans. Signal Proc.*, vol. 55, no. 6, pp. 2542–2556, 2007.

[60] M. A. Poletti, T. D. Abhayapala, and P. N. Samarasinghe, "Interior and exterior sound field control using two dimensional higher-order variable-directivity sources," *J. Acoust. Soc. Am.*, vol. 131, no. 5, pp. 3814–3823, 2012.

[61] B. Rafaely and D. Khaykin, "Optimal Model-Based Beamforming and Independent Steering for Spherical Loudspeaker Arrays," *IEEE Trans. Audio, Speech Lang. Proc.*, vol. 19, no. 7, pp. 2234–2238, 2011.

[62] H. F. Olson, "Gradient Loudspeakers," *J. Audio Eng. Soc.*, vol. 21, no. 2, pp. 86–93, 1973.

[63] T. D. Abhayapala and D. B. Ward, "Theory and design of high order sound field microphones using spherical microphone array," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2002, vol. 2, pp. 1949 –1952.

[64] B. N. Gover, J. G. Ryan, and M. R. Stinson, "Microphone array measurement system for analysis of directional and spatial variations of sound fields.," *J. Acoust. Soc. Am.*, vol. 112, no. 5 Pt 1, pp. 1980–1991, 2002.

[65] B. Rafaely, "The spherical-shell microphone array," *IEEE Trans. Audio, Speech Lang. Proc.*, vol. 16, no. 4, pp. 740–747, 2008.

[66] A. Laborie, R. Bruno, and S. Montoya, "A new comprehensive approach of surround sound recording," in *AES 114th Convention*, 2003.

[67] Z. Li and R. Duraiswami, "Flexible and Optimal Design of Spherical Microphone Arrays for Beamforming," *IEEE Trans. Audio, Speech Lang. Proc.*, vol. 15, no. 2, pp. 702–714, 2007.

[68] E. Tiana-Roig, F. Jacobsen, and E. F. Grande, "Beamforming with a circular microphone array for localization of environmental noise sources.," *J. Acoust. Soc. Am.*, vol. 128, no. 6, pp. 3535–3542, 2010.

[69] H. Teutsch and W. Kellerman, "Acoustic source detection and localization based on wavefield decomposition using circular microphone arrays," *J. Acoust. Soc. Am.*, vol. 120, no. 5, pp. 2724–2736, 2006.

[70] D. N. Zotkin, R. Duraiswami, and N. A. Gumerov, "Plane-wave decomposition of acoustical scenes via spherical and cylindrical microphone arrays," *IEEE Trans. Audio, Speech Lang. Proc.*, vol. 18, no. 1, pp. 2–16, 2010.

[71] J. Meyer, "Beamforming for a circular microphone array mounted on spherically shaped objects," *J. Acoust. Soc. Am.*, vol. 109, no. 1, pp. 185–193, 2001.

[72] J. Meyer and T. Agnello, "Spherical microphone array for spatial sound recording," in *AES 115th Convention*, 2003.

[73] A. Parthy, N. Epain, A. van Schaik, and C. T. Jin, "Comparison of the measured and theoretical performance of a broadband circular microphone array," *J. Acoust. Soc. Am.*, vol. 130, no. 6, pp. 3827–3837, 2011.

[74] N. Epain and J. Daniel, "Improving spherical microphone arrays," in *AES 124th Convention*, 2008.

[75] J. Meyer and G. W. Elko, "Spherical microphone arrays for 3D sound recording," in *Audio signal processing for next-generation multimedia communications systems*, Amsterdam, The Netherlands: Kluwer Academic Publishers, 2004, pp. 67–90.

[76] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 1, pp. 135–143, 2005.

[77] M. Park and B. Rafaely, "Sound-field analysis by plane-wave decomposition using spherical microphone array," *J. Acoust. Soc. Am.*, vol. 118, no. 5, pp. 3094–3103, 2005.

[78] D. B. Ward and T. D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 6, pp. 697 – 707, 2001.

[79] D. B. Ward, R. A. Kennedy, and R. C. Williamson, "Theory and design of broadband sensor arrays with frequency invariant far-field beam patterns," *J. Acoust. Soc. Am.*, vol. 97, no. 2, pp. 1023–1034, 1995.

[80] H. Kuttruff, *Room Acoustics*. London: Applied science publishers Ltd, 1973.

[81] "Playrec: Multi-channel Matlab Audio." [Online]. Available: http://www.playrec.co.uk. [Accessed: 21-Oct-2014].

[82] N. Aoshima, "Computer-generated pulse signal applied for sound measurement," *J. Acoust. Soc. Am.*, vol. 69, no. 5, pp. 1484–1488, 1981.

[83] T. Betlehem and T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," *J. Acoust. Soc. Am.*, vol. 117, no. 4, pp. 2100–2111, 2005.

[84] L. Bianchi, F. Antonacci, A. Canclini, A. Sarti, and S. Tubaro, "A psychoacoustic-based analysis of the impact of pre-echoes and post-echoes in soundfield rendering applications," in *International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2012.

[85] Y. Yai, S. Miyabe, H. Saruwatari, K. Shikano, and Y. Tatekura, "Rapid compensation of temperature fluctuation effect for multichannel sound field reproduction system," *Ieice Trans. Fundam.*, vol. E91-A, no. 6, pp. 1329–1336, 2008.

[86] P.-A. Gauthier and A. Berry, "Adaptive wave field synthesis with independent radiation mode control for active sound field reproduction: Theory," *J. Acoust. Soc. Am.*, vol. 119, no. 5, p. 2721, 2006.

[87] M. Schneider and W. Kellermann, "Adaptive listening room equalization using a scalable filtering structure in the wave domain," in *Proc. Itl. Workshop on Acoust. Signal Enhancement, IWAENC*, 2012.

[88] A. Canclini, P. Annibale, F. Antonacci, A. Sarti, R. Rabenstein, and S. Tubaro, "A methodology for evaluating the accuracy of wave field rendering techniques," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, 2011, pp. 69–72.

FIGURE CAPTIONS

Figure 1: Third order loudspeaker unit

Figure 2: Higher order microphone

Figure 3: Reproduction room with surround sound layout

Figure 4: Nett impulse responses at several positions in the reproduction zone, comparing AC-SFR using (a),(b),(c) zeroth order secondary sources and (d),(e),(f) third order secondary sources with second order Ambisonics (max-r_E) for a virtual source angle of 0 degrees.

Figure 5: Nett impulse responses at several positions in the reproduction zone, comparing AC-SFR using (a),(b),(c) zeroth order secondary sources and (d),(e),(f) third order secondary sources with second order Ambisonics (max-r_E) for a virtual source angle of 36 degrees.

Figure 6: Nett frequency responses at several positions in the reproduction zone, comparing AC-SFR using (a),(b),(c) zeroth order secondary sources and (d),(e),(f) third order secondary sources with second order Ambisonics (max-r_E) for a virtual source angle of 36 degrees.

Figure 7: Beamformer SRP at the centre of the sound control zone in various frequency bands for (a) AC-SFR with zeroth order sources, (b) AC-SFR with third order sources and (c) Second order Ambisonics. Plots are shown for a virtual source angle of 36 degrees. The directions to the secondary sources are also marked (--).

Figure 8: Beamformer SRP of the pre-reverberation in several frequency bands for a virtual source at 36 degrees, midway between two loudspeakers, for a room with 50mm ceiling absorption.

Figure 9: Relative pre-reverberation levels versus frequency