

Sound field Synthesis with Distributed Modal Constraints

Dylan Menzies¹

¹Department of Media Technology, De Montfort University, LE1 9BH, Leicester, UK rdmg@dmu.ac.uk

Abstract

A method is introduced for calculating driving functions for loudspeaker arrays used to control sound fields. This is based on the simultaneous solution of multiple modal constraints whose centres are distributed in space. Several test examples are described. Subjects considered include the control of the complete interior of convex and concave array boundaries, the control of subregions, loudspeaker distributions, open boundaries, independent regions, point source targets, and encoding and decoding.

1 Introduction

Sound field synthesis is the reproduction of any sound field over a region of space using fixed loudspeakers. In this article the surrounding space is considered to be anechoic. There is no scattering of sound back into the region. The control is purely feedforward, with no feedback from sensors. Sound field synthesis can be expressed exactly using a continuum of sources. In practice only a finite number of drivers can be used, and reproduction error is unavoidable. A variety of methods have been developed with different error and calculation properties. These were reviewed and compared together at a recent conference [1] to highlight some possible areas for development. Some proposals were discussed, including one referred to as Distributed Modal Constraints (DMC). Parts of this work are expanded here with the focus on DMC.

The next section reviews existing methods for sound field synthesis using fixed drivers. Usage and performance characteristics are highlighted rather than implementation details. In the following section the development and implementation of DMC is described. Simulations are then presented covering a range of synthesis tasks. Comparisons

are made with existing systems in these cases. Finally the problem is considered of how to encode and decode DMC so that it can be rendered on different arrays.

2 Sound field systems review

2.1 Ambisonics

In its original form the Ambisonics system consists of a method of approximately encoding 3-dimensional sound fields with 4 channels and a method for decoding these channels to surrounding loudspeaker arrays in order to reproduce the encoded sound field in the interior [2]. The encoding approximates the directivity of sound incident at a point in space, and the loudspeaker reproduction approximately reconstructs this directivity at the centre of the array. High-order Ambisonics (HOA), [3], is an extension of Ambisonics based on the Fourier-Bessel expansion (FBE), allowing more accurate sound field encodings and reproductions using more channels and speakers. Increasing the encoding order increases the angular resolution and also allows a sound field to be accurately represented over a wider region at each frequency. For plane waves, and most naturally occurring fields that are well separated from sources, the spherical region $r < N/k$ is accurately represented by using an FBE to order N , for radius r and wavenumber k [4, 5]. Fields for which this is not true are referred to in this article as *badly behaved*. A simple example is defined by giving non-zero power to a single mode with order greater than kr . A more natural example is given by a monopole. As the region moves closer to the monopole the error progressively worsens from that given by the $r < N/k$ criterion, [6]. In general for $r > N/k$ the error is non-uniformly distributed [7].

In the decoding step the sum of the FBEs of the

contributing loudspeaker sources is matched to the target FBE [3], also called mode-matching. The first part of Section 3.1 gives details of the decoding formulation. If there are sufficiently many, well distributed loudspeakers, then the sound field represented by the FBE can be accurately reproduced, [4]. Spherical and circular arrays of evenly distributed drivers are most commonly considered, as these match the geometry of the reconstructed region, and have the simplest driving functions.

The driver spacing required to completely control the interior of a spherical or circular array up to a given frequency can be estimated as follows. For the 2D case the number of drivers is $L = 2N + 1$, [3]. The perimeter is $P = 2\pi r$. The driver spacing $\approx P/L \approx P/(2N) = P/(2kr) = (P\lambda)/(4\pi r) = \lambda/2$. Similarly for 3D, an even distribution on a sphere can be locally approximated with area per source $\rho = (\delta\lambda)^2$, where $\delta\lambda$ is the driver spacing. The number of sources is $L = (N + 1)^2$, [3]. The total boundary area $A = 4\pi r^2 = L\rho \approx N^2\rho = (2\pi r/\lambda)^2(\delta\lambda)^2$, from which $\delta = \sqrt{1/\pi} \approx 0.56$. In both 2D and 3D cases the value of δ is similar to that expected from cartesian sampling in a line or plane.

Decoding may be applied to any shape of driver boundary, or even any distribution of drivers. The reproduction region remains spherical at each frequency, as it is determined by the FBE. It is not possible to control the whole interior of a non-spherical boundary as the FBE of each source about the array centre is valid only as far as the radius extending from the centre to the source. If the FBE were valid beyond the source it would include the source, which is not possible because any FBE field is sourceless. Beyond the source the FBE differs to that of a real source, so the fields that are represented by mode matching are different to those actually generated by the sources, and the field is not correct over the whole region.

To illustrate the last point Fig. 1 shows a $(20\lambda, 5\lambda)$ size simulated rectangular array, with source spacing $\lambda/2$. An accurate reproduction is shown of a 2D $N = 20$ FBE of a 2D plane wave travelling from the top. The first plot shows the real part of the field, the second shows the absolute relative error $|(p - \hat{p})/\hat{p}|$, where p is the synthesized pressure and \hat{p} is the target pressure at a point. The same keys are used for comparisons across all the figures in this article. The expansion

order limit based on a limited radius equal to half the width is $N = (2.5)(2\pi) \approx 15$. Increasing N only increases the region slightly up to the limit shown for $N = 20$. The error transition at the FBE edge is swift, as expected. Further increase in N actually causes the error in the central disk to increase as shown in Fig. 2. The maximum size of the reconstructed region cannot be improved by change of order, source spacing or regularization. Fig 3 shows schematically the largest FBE area that can represent the closest source.

From a computational viewpoint, HOA consists of two steps, encoding and decoding. The decoding is applied only once to the sum of all encoded sources. Each encoded source requires N filter calculations. Decoding requires a separate filter for each component of the encoding, $(N + 1)^2$ components in 3D and $2N + 1$ in 2D. A feature of HOA is that an encoding can be rotated prior to decoding, without using filters. This is useful in virtual reality applications where the head rotates relative to a complex background environment [8].

Realistic drivers are complex sources, rather than point sources, and can be approximated as multipole expansions. Driving functions for fixed directivity cardioid drivers have been derived using mode matching constraints [9]. It was also shown how separate dipole drivers can be controlled to minimize external radiation, as would be expected from the Kirchhoff-Helmholtz integral.

HOA encodings have been constructed for multiple listening zones [10, 11], in which at least two regions appear to have independent sound fields, although they are actually part of the same enclosing sound field. If decoded with mode-matching, the same limitations apply as before.

2.2 Wave Field Synthesis

The original formulation of Wave Field Synthesis (WFS) is based on Rayleigh's integrals, which can be derived from the Kirchhoff-Helmholtz integral. A sound field is constructed in a half-space from a continuum of monopole sources on the boundary [12, 13]. This half-space is restricted to free-fields that can be composed of plane waves travelling away from the boundary. The driving functions for plane and point target sources are computationally efficient consisting of a single pre-filter and a delay depending on driver position. Additional weights

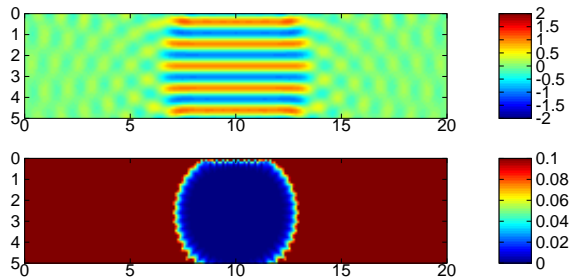


Figure 1: An FBE of a plane wave with $N = 20$ decoded for a rectangular array.

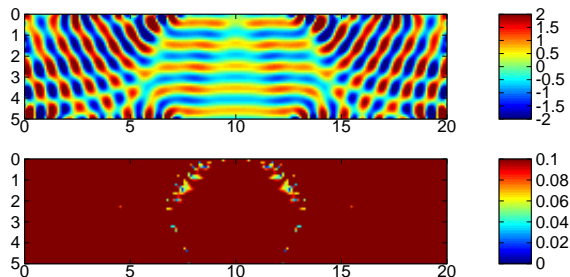


Figure 2: An FBE of a plane wave with $N = 26$ decoded for a rectangular array.

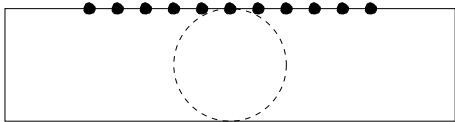


Figure 3: Limit of representation of the closest source by an FBE at the centre of a rectangular array.

can be introduced to minimize truncation effects. Approximate control with a linear rather than planar source boundary is possible using a modified pre-filter. For a real array the driving function must be sampled at intervals. Where the intervals are equal, this causes error that increases rapidly for $\delta > 1/2$, as for HOA. However in this case the error is usually distributed in space around the array [7].

Linear array WFS systems can be extended by combining several line arrays to make a convex enclosure, or more generally placing drivers along a curved boundary. These geometries no longer satisfy the conditions of Rayleigh's integral, however surprisingly low error is obtained by using the same

plane wave driving function as before with an added window term that switches the driver off when the interior surface normal is opposed to the target plane wave direction. The error introduced by joining two line arrays is referred to as *corner effect* in [14], and in general any non-linear or non-planar boundary shape creates additional error, although this is not fully understood [15]. The above extension for non-linear boundaries does not work for concave boundaries. The region around a single concave point cannot be covered by a locally generated plane wave field for some angles. A concave point is any point where there is an external centre of curvature, or it is where two line segments join with an interior angle greater than 180 degrees. Concave boundaries are potentially useful for extending sound fields across connecting rooms.

Interior point sources, in other words those located in the reproduction area, can be approximated by *focused sources* [14], for which the driving functions are similar to those for the exterior sources. Focused sources are similar to a point source in a half-space away from the array, but the error increases significantly inside one wavelength from the source centre, where the focused source is limited and the point source is unlimited. With access to all plane wave directions it is possible to better approximate point sources in a half space within one wavelength, but this requires increased driving power and higher resolution. In HOA both approximations of the point source can be produced directly [16, 17].

Focused sources have unusual error properties in the context of WFS. The drivers can be spatially subsampled, meaning $\delta > 1/2$, and still yield a focused source that is accurate around its centre [18]. This has recently been exploited in a proposal for controlling local regions, *local WFS* [19]. A virtual driver array composed of focused sources is controlled by the real array, which contains the virtual array. The virtual array is driven to control the target required using WFS, and this then determines how the real array is driven. This enables synthesis of convex subregions which is efficient computationally and in terms of numbers of drivers used. This is related to the situation in HOA where $\delta > 1/2$. It is not clear from [19] how local WFS would perform for elongated boundaries.

A WFS enclosure can reproduce approximate plane waves at all angles, so it can approximate

any free field, including those represented by HOA encodings. This offers the advantage over HOA decoding of reproducing the HOA encoding over the entire interior of a WFS boundary of any shape. The total cost of such a WFS decoding is comparable with HOA decoding as filters result from the evaluation of the mode velocities at the boundary.

There is no direct way to produce a multi-zone sound field using WFS, although the WFS decoding approach just described could be applied to a multi-zone HOA encoding.

2.3 Sampled Simple Source

The Kirchhoff-Helmholtz integral provides a construction for any interior free-field defined in any closed boundary in terms of monopole and dipole sources on the boundary. This can be manipulated to provide a construction just using monopoles, known as the simple source formulation [5]. This can be approximated for a practical realization by sampling the driving function and is referred to as Sampled Simple Source (SSS).

For circular arrays in 2D and spherical arrays in 3D, simple source driving functions can be found by applying simple functions to the HOA encoding [4]. These driving functions coincide with mode matched decodings for some symmetric arrays including the 2D polygons and 3D platonic solids. As the number of sources increase and they become more closely spaced, SSS and HOA solutions converge. Because the simple source decoding is simpler and has other attractions it is sometimes preferred over mode matching, and has become part of the overall HOA methodology.

For general boundaries the simple source driving functions can be found by numerical solution. The direct approach according to the formulation of the simple source representation is equivalent to solving a scattering problem [20]. This can be implemented using a boundary element approach, such as the fast multipole method [21], and solving it for the velocity of the scattered field on the boundary. For convex boundaries, SSS and WFS converge towards high frequency [20].

2.4 Pressure Constraints

Driving functions can be found numerically by satisfying pressure constraints at a set of points

[22, 20, 23], abbreviated here as the PC method. In some cases accurate solutions can be found, however in general poor conditioning and numerical difficulties can degrade solutions. The solution sought is not exactly well defined as there always exist non-trivial fields at certain eigenfrequencies, for which the pressure constraints are all zero [5]. When the constraints are on a simple shell the eigenfrequencies will be in the range of interest.

Some active control systems use pressure microphones to adapt the driving functions to compensate for room acoustics. The initial driving functions are for an anechoic environment using a feed-forward synthesis procedure. There are two examples of this approach based on WFS [23, 24]. The microphones alone cannot identify the field. Stability is provided by the free field WFS driving functions, which are known.

2.5 Summary

The following list summarizes features that may be desired in spatial synthesis and how these are met by the methods discussed. The particular combination of features favoured in a given application varies, and so no approach is always preferable.

- Full interior control for general convex boundaries. Possible with WFS and PC but not HOA.
- Subregion control, with subsampled boundary. This is inherent with HOA, and possible with PC in some cases. Local WFS looks promising but it is not clear, for example, how this would perform for elongated regions in 3D.
- Full interior and subregion control for general concave boundaries. This has not been demonstrated using any of the methods. SSS can be used for concave boundaries but not subregions, and in any case calculation is relatively difficult.
- Control in multiple regions. This has been demonstrated by constructing an HOA encoding with constrained subregions, but not for general boundary and subregion geometry.
- Low computational cost. Per source WFS has lowest costs, consisting of interpolated delays

and one filter. HOA encoding requires more filters per source, and then decoding.

- Sound field encoding method for storing a complex scene, with low complexity rotation. HOA encodings are an example of this. WFS decoding for HOA encodings is possible, but the advantage is not clear.

It is evident that two possible areas for development are subregions and concave boundaries. These provide the initial motivations for developing DMC.

3 Distributed Modal Constraints

From an intuitive viewpoint, PC has only indirect knowledge of the wave equation, through the sampling of each driving function by the constraints. On the other hand in HOA knowledge of the wave equation is implicit in the representation of the field by modes satisfying the wave equation. The $N = kr$ criterion implies knowledge of the sound field in a non-vanishing sphere (3D) or disc (2D) when $N \geq 1$. A pressure constraint is equivalent to $N = 0$ and has no spatial extent of such a kind. This suggests that the effectiveness of PC can be improved by replacing each pressure constraint by an HOA type modal constraint, in other words Distributed Modal Constraints (DMC).

DMC is related to multi-zone HOA, but can be applied to a wider range of problems. The constraints are applied directly in the calculation of the driving functions, allowing FBE regions to be placed anywhere within the boundary. The size of each region given by the radius $r = N/k$ determines how close the region centre can be to the boundary. Regions that are separated can either sample the same underlying target field or else sample different target fields, which is the multi-zone case. Overlapping FBE regions are also considered, in which case the constraints must sample from the same well behaved underlying field. The smallest non-trivial FBE region is 1st order. Overlapping 1st order regions therefore provide the highest resolution for defining a general shape of a controlled region.

Fig 4 illustrates three cases for FBE region placement, overlapped, touching and separated.

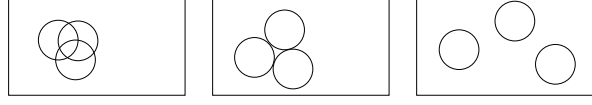


Figure 4: Modal regions, overlapped, touching and separated.

3.1 Formulation

The modal constraint of an Ambisonic decoder for freely located loudspeaker sources, in the frequency domain, can be written as

$$c_n = \sum S_{jn}(\mathbf{r}_{s_j} - \mathbf{r}_c)s_j = \sum M_{nj}s_j, \quad (1)$$

where c_n are the coefficients of the constrained modes and s_j are the loudspeaker source complex weights. Formulae here use a single modal index n that can apply to either the 2D or 3D case. $S_{jn}(\mathbf{r}_{s_j} - \mathbf{r}_c)$ is the n th component of the FBE decomposition of the j th unweighted source with location $\mathbf{r}_{s_j} - \mathbf{r}_c$ relative to the FBE centre. \mathbf{r}_{s_j} and \mathbf{r}_c are the locations of the sources and the FBE centre respectively. In the simplest case $S_{jn} = S_n$ describes identical monopoles, but any function can be used, for instance to model a speaker more accurately, or incorporate room reflections. The matrix M is defined for convenience. The decoder is solved by finding the pseudo inverse of M , M^+ which aims to satisfy

$$s_j = \sum M_{jn}^+ c_n. \quad (2)$$

Regularization of the inversion can be used to improve numerical stability and limit source coefficient magnitude and in return for some increase in solution error. The simplest form of Tikhonov regularization is used in this article, so that M^+ becomes the regularized pseudo inverse $(M^T M + \alpha^2 I)^{-1} M^T$, with parameter α . This regularization yields smoothly varying source coefficients as frequency is varied. Regularization by zeroing singular values, such as that built into Matlab's *pinv* function, is not appropriate because the source coefficients jump at frequencies where a singular value is lost, and efficient filters cannot be formed.

α is dimensionless as both s and c measure pressure, but in different geometric contexts. However choice of α does depend on problem geometry and parameters. If the density of sources is doubled

then the source signals are about half as big as the original. M increases in dimension, and the singular values all increase, by nearly the same factor. Then a larger value of α achieves a similar result. The fraction relative to the largest singular value σ_{max}/α gives a measure of regularization that is independent of the density change, and is also the condition number of the regularized system matrix. In the simulations the values of α are quoted for completeness, and their choice is explained. In most cases the reproductions are very similar for a wide range of values.

To extend to constraints with multiple FBE centres, l is used to index them. Then the total set of constraints is written

$$c_{ln} = \sum S_{jn}(\mathbf{r}_{s_j} - \mathbf{r}_{c_l}) s_j = \sum M_{lnj} s_j \quad . \quad (3)$$

The l and n indices can be combined into one index \bar{n} with NL values, where N is the number of modes and L is the number of FBE centres. Then c_l can be written in terms of \bar{n} , and the constraint equation can be expressed as a single matrix equation,

$$\bar{c} = \bar{M} s \quad , \quad (4)$$

where \bar{M} is indexed by \bar{n} and j , and \bar{c} by \bar{n} . The pseudo inverse \bar{M}^+ of the extended matrix \bar{M} attempts to satisfy

$$s = \bar{M}^+ \bar{c} \quad . \quad (5)$$

Regularization will again have the effect of limiting source energy for some reduction in solution accuracy. The formulation has made a shift away from the single centre view of Ambisonics, including multi-zone methods with a single Ambisonic encoding.

In the simulations considered here the boundaries are transparent, with point source drivers, and there is no room acoustic contribution. The resulting basis functions and expansions for 2D and 3D cases are now described. Positive and negative frequency versions are provided here as a useful reference.

In 2D, or equivalently 3D with line sources, FBE basis functions $J_n(kr)e^{in\phi}$ are used. Then with the positive frequency convention, which is most common in Ambisonic literature, the loudspeaker source function $H_0^{(2)}(kr)$, where r is the radius from

the source centre, has components

$$S_n(\mathbf{r}) = H_n^{(2)}(kr)e^{-in\phi} = H_n^{(2)}(kr)((x - iy)/r)^n \quad , \quad (6)$$

where \mathbf{r} is the location of the source relative to the expansion centre. $\mathbf{r} = (r, \phi)$ in polar coordinates and (x, y) in cartesian [25]. The negative frequency convention loudspeaker source $H_0^{(1)}(kr)$ has components $S_n(\mathbf{r}) = H_n^{(1)}(kr)e^{-in\phi}$. Note that scaling the loudspeaker source functions has no effect on the solution field. The components c_{ln} for a simple source target field are

$$c_{ln} = S_n(\mathbf{r}_s - \mathbf{r}_{c_l}) \quad , \quad (7)$$

where \mathbf{r}_s is the location of the target source. The dependence on \mathbf{r}_{c_l} ensures the multiple expansions consistently sample from the same target field.

The other type of target field considered here is a plane wave. In the positive frequency convention $e^{-i\mathbf{k}\cdot\mathbf{r}}$ represents a wave traveling in the direction of \mathbf{k} . The components at the expansion centres are

$$c_{ln} = (-i)^n e^{-in\phi_{\mathbf{k}}} e^{-i\mathbf{k}\cdot\mathbf{r}_{c_l}} \quad , \quad (8)$$

where $\mathbf{k} = (k, \phi_{\mathbf{k}})$ in polar coordinates. The last factor is a phase adjustment due to the location of each expansion centre. With the negative frequency convention a wave traveling with direction \mathbf{k} is instead given by $e^{i\mathbf{k}\cdot\mathbf{r}}$, and has components $i^n e^{-in\phi_{\mathbf{k}}} e^{i\mathbf{k}\cdot\mathbf{r}_{c_l}}$.

In 3D the N3D harmonics, $Y_{mn}(\hat{\mathbf{r}})$ can be used [3]. These are a natural extension of the functions used in low order Ambisonics. The basis functions are $i^m j_m(kr)Y_{mn}(\hat{\mathbf{r}})$. The literature is inconsistent in the lettering and naming of the indices for harmonics. The first index m is called here the *order*, with range from 0 to N , the maximum order. This is the more usual naming but is contrary to that used in [3]. The equation $N = kr$ applies in both the 2D and 3D cases. In the positive frequency convention a simple source $\frac{e^{-ikr}}{kr}$ located at position \mathbf{r} relative to the expansion centre has components

$$S_{mn}(\mathbf{r}) = i^{-(1+m)} h_m^{(2)}(kr) Y_{mn}(\hat{\mathbf{r}}) \quad , \quad (9)$$

where the index n is replaced by the two indices m, n . In the negative frequency convention for a simple source $\frac{e^{ikr}}{kr}$ the components are $i^{(1-m)} h_m^{(1)}(kr) Y_{mn}(\hat{\mathbf{r}})$.

As for the 2D case, $e^{-i\mathbf{k}\cdot\mathbf{r}}$ in the positive frequency convention represents a plane wave traveling in the direction of \mathbf{k} , and has components

$$c_{lmn} = Y_{mn}(-\hat{\mathbf{k}})e^{-i\mathbf{k}\cdot\mathbf{r}_{cl}} \quad , \quad (10)$$

again with a phase factor depending on the expansion centre. In the negative frequency case $e^{i\mathbf{k}\cdot\mathbf{r}}$ is a plane wave traveling along \mathbf{k} , with components $Y_{mn}(\hat{\mathbf{k}})e^{i\mathbf{k}\cdot\mathbf{r}_{cl}}$.

For convex enclosures with interior corner angles less than 180° , transparent boundaries give the same interior field as absorbing boundaries. For concave boundaries this is not so, and for non-transparent walls diffraction must be considered. In principle, scattering, diffraction and source directivity could be included by simulating the response to each driver using acoustic simulation software or measurement, then encoding in the expansion components $S_{jn}(\mathbf{r}_{sj} - \mathbf{r}_{cl})$. This is not pursued here.

3.2 Simulations

Some tests of DMC are now presented using a variety of target sound fields and boundaries. In each case plots have been produced showing the simulated fields and the absolute relative error. To simplify the investigation the boundaries are considered to be transparent, with 2D point source drivers. From previous discussion, where WFS or HOA can be used to control a whole interior, the driver spacing is at most $\delta = 1/2$. In the following simulations using DMC $\delta = 1/2$ is chosen initially, and then varied.

3.2.1 Touching and overlapping modal regions

In the first simulation the example in Fig. 1 is revisited, this time with 4 neighboring modal constraints each of order $N = 15$. The results are shown in Fig. 5. The parameters shown include the number of centres L , the number of loudspeaker sources J , the source spacing as a fraction of a wavelength δ , the regularization parameter α , the maximum absolute value of the source coefficients $\max(s)$. The maximum absolute value of the field $\max(p)$ will be introduced later. The plane wave and source fields are as described in the previous section. As p is unbounded towards each source,

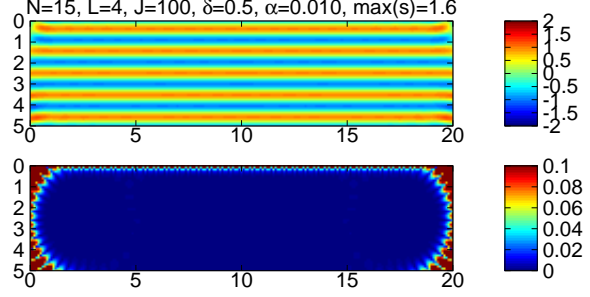


Figure 5: Reproduction of plane wave from the top, with 4 touching modal constraints.

$\max(p)$ is considered everywhere in the interior except points that closer than $\delta\lambda$ from the boundary. Both $\max(s)$ and $\max(p)$ can be used to judge how practical a solution is. $\max(s)$ depends on source spacing, but this is not varied much in the following simulations.

The value of α is initially chosen to be as small as possible while keeping $\max(s)$ comparable to the target field. Further decrease reduces reproduction error at the expense of excessive source strength.

With 4 centres the error is low across the entire region, with obvious significant error outside the end modal regions. The gaps between modal regions have been bridged, although increased error in those parts is still apparent.

For the angled plane wave shown in Fig. 6, again 4 centres, the error is significantly higher. Using 100 equally spaced regions the error is lowered across the whole target region, Fig. 7. The high degree of degeneracy created by the modal overlap does not affect the solution adversely, and $\max(s)$ is only slightly increased.

Reducing source density slightly disrupts the solution, as shown in Fig. 8 where $\delta = 0.6$. Increasing the density reduces the error slightly. This supports the original assumption that a spacing of $\lambda/2$ is a good choice.

3.2.2 Maximum resolution coverage

In the previous examples the corner areas are unconstrained, and consequently show raised error. These areas can be covered with additional smaller constraint regions. In general, target regions can be covered with the greatest spatial accuracy using $N = 1$ modes. Fig. 9 shows the result for a 500

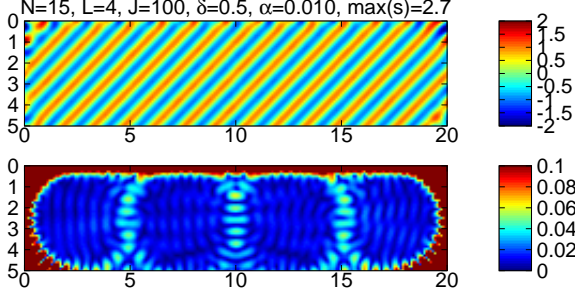


Figure 6: Reproduction of plane wave from the top-left, with 4 touching modal constraints.

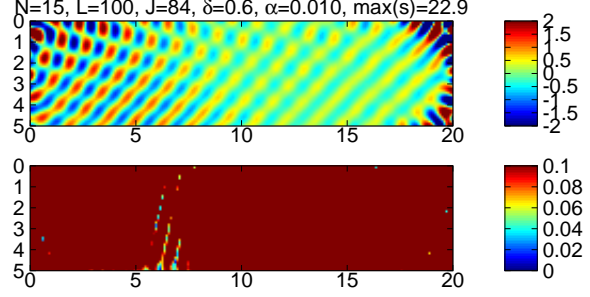


Figure 8: Reproduction of plane wave from the top-left, with 100 overlapping modal constraints. Source spacing has been increased slightly.

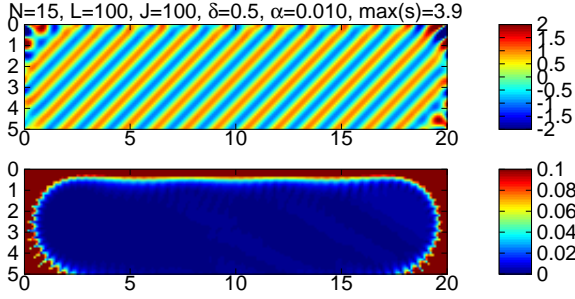


Figure 7: Reproduction of plane wave from the top-left, with 100 overlapping modal constraints.

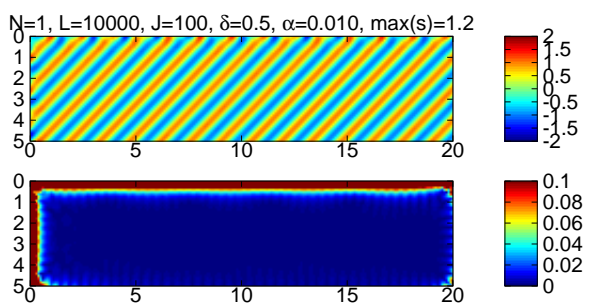


Figure 9: Reproduction of plane wave from the top-left, with 10000 interior modal constraints.

by 20 array of modal regions, with a $\lambda/3$ margin separating the boundary from the centres. Compared with the previous case lower error is achieved with lower $\max(s)$. $N = 1$ constraints are equivalent to combined pressure and gradient constraints. Such constraints defined continuously on a closed boundary define the interior sound field uniquely, according to the Kirchhoff-Helmholtz integral, so it would be expected that the interior constraints are to some degree redundant. Fig. 10 shows a simulation with the same number of constraints located on the boundary of the constrained region. The solution is worse than with interior constraints, and changing the number of constraints does not significantly improve the solution. As the number of sources increase, both types of solution converge to the target.

3.2.3 Concave boundaries

Consider an L-shaped boundary, with FBE centres placed on an L-shaped line in the interior. Fig. 11

shows a reproduced wave moving from the top left. The reproduction problem is similar to that for a rectangular interior. Little power is needed around the concave part. In Fig. 12 the wave moves in the other direction, sourced mostly from the concave part. The error and $\max(s)$ are increased. Increasing the source density with $\delta = 0.4$ in Fig. 13 produces low error and power solution. The remaining error is seen to focus around the concave point. This suggests that an optimum distribution of sources will be slightly more dense at points of increased curvature, and smoothing out sharp corners will help to improve the solution for a given number of sources.

3.2.4 Source distribution and control of subregions

Two questions arise. If the whole interior is controlled, how much freedom is there to redistribute the sources? If the constrained region is a subset

$N=1, L=10000, J=100, \delta=0.5, \alpha=0.010, \max(s)=1.2$

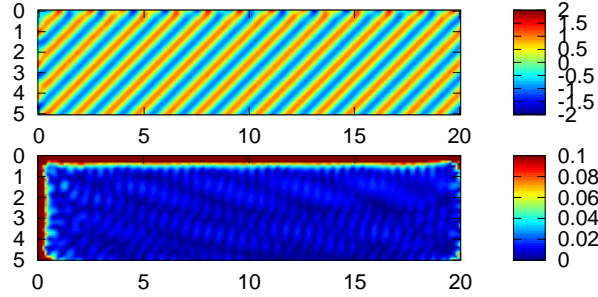


Figure 10: Reproduction of plane wave from the top-left, with 10000 boundary modal constraints .

of the interior can the number of sources be reduced, and how? For the first question it is clear that for a long rectangular interior, moving sources to the ends and thinning out the sides will not be acceptable since the end contribution is small. So in general only a small amount of redistribution is possible.

For the second question again consider a long rectangular interior, with a narrow central target region. If the narrow region were controlled by a second boundary of sources that directly enclosed it, the required source spacing is $\lambda/2$, from the first question. The same number of sources are required on the outer boundary, so the spacing required there is the same, Fig. 14. In other words the outer spacing cannot be increased to control the narrow region. The number of sources required is $J = 2P/\lambda$ where P is now redefined as the perimeter of the controlled region. $J \sim l$ where l is the length of the region. To achieve control of the same region using an Ambisonic array, as shown in Fig. 15, would require more speakers, but still $J \sim l$.

In 3D the picture is different. Consider a long tubular boundary. The interior can be controlled with sources on the boundary with spacing $\lambda/2$. The same number of sources are needed to control this region from an outer boundary, as shown in Fig. 14, by comparison with 2D HOA. The total number of sources is $J = 4A/\lambda^2 \sim l$ where A is the surface area of the controlled region, and l the length of the tube. If the outer boundary is further distorted to the shape of a sphere, Fig. 15, further simulation shows that the number of sources required is the same, provided they are redistributed

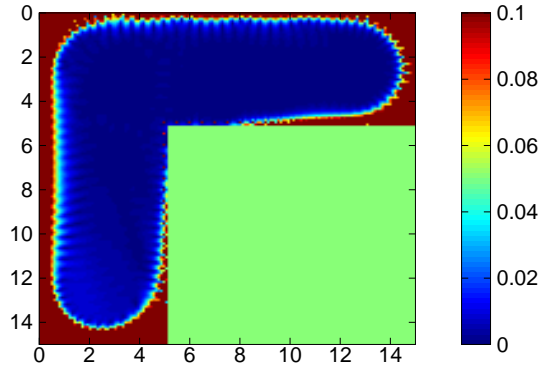
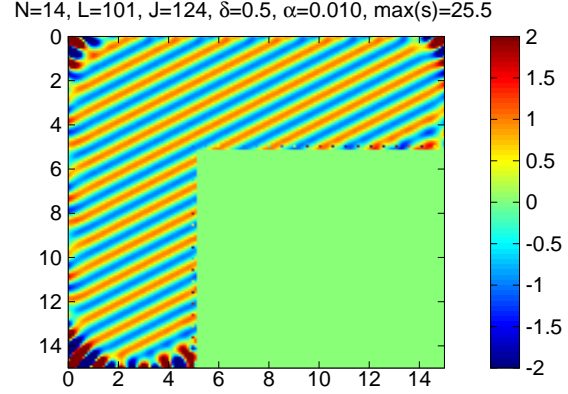
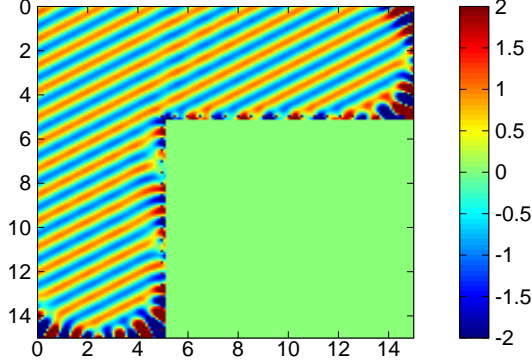


Figure 11: L-shaped boundary and control region. Wave from top left.

so that those in contact with the tube are spaced by $\lambda/2$, with spacing increasing with distance from the tube. The cross sections look like Fig. 14. On the other hand the whole sphere interior could be controlled using 3D HOA with $J \sim l^2$, which is a factor of l more expensive than control using DMC. This could be valuable in situations where the listeners are constrained to move along lines, for instance gangways in museums, and in spherical projection theatres.

Optimal source configuration for any boundary and controlled subregion is not investigated in detail here, however simulations such as the above example indicate that for any region with perime-

N=14, L=101, J=124, $\delta=0.5$, $\alpha=0.010$, $\max(s)=10.0$



N=14, L=101, J=154, $\delta=0.4$, $\alpha=0.010$, $\max(s)=2.2$

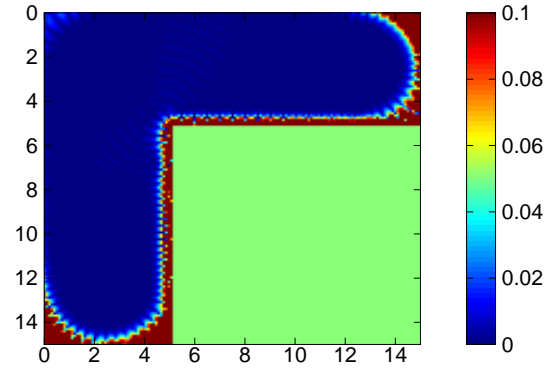
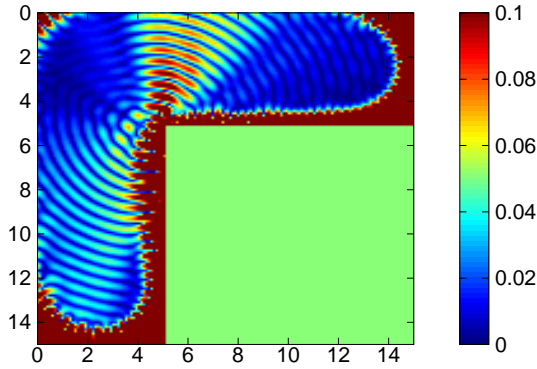
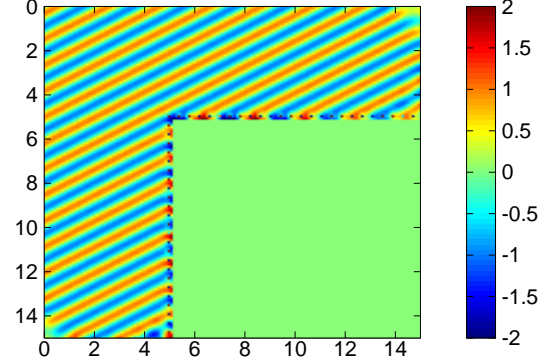


Figure 12: L-shaped boundary and control region. Wave from bottom right.

Figure 13: L-shaped boundary and control region. Wave from bottom right. Increased source density.

ter P or area A , and an outer boundary, a boundary configuration of $2P/\lambda$ or $4A/\lambda^2$ sources can be found that controls the region. A concave region is a subset of a convex region with lower P or A , and so requires less sources than implied by the P or A of the concave region. However not all concave boundary fields are reproducible using enclosing convex boundaries. The question is revisited in the later section on encoding.

3.2.5 Open boundaries

In the previous examples most of the power comes from sources on parts of the boundary where the wave travels from. If the opposite sources are re-

moved, it is still possible to form good waves over a limited range of directions. Fig. 16 shows the reproduction of a plane wave using just a single line array at the top, which is the same configuration as a linear array WFS system. The controlled region is a bar from location $(5, 5)$ to $(15, 5)$. In the example the error is low in the control region and also in a region beyond this. Additional constraints can be used to improve the reproduction in a triangular region for a range of plane wave angles. As the angle of wave direction moves closer to the direction of the array, reproduction quality reduces, and an inward moving wave is impossible. $\max(s)$ can be reduced to 3 with $\alpha = 0.1$, for a small increase in

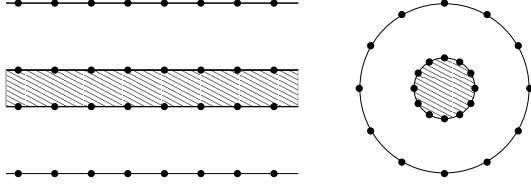


Figure 14: Source on inner and outer boundaries in 2D (left only) and 3D (both, cross section). The controlled region is shaded.

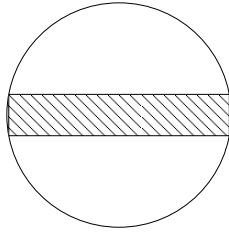


Figure 15: Rectangle or tube enclosed by a circular or spherical array. The controlled region is shaded.

error in the control region. Reducing δ to 0.4 does not have a significant effect.

Another example of an open boundary is a semi-circle (2D) or hemisphere (3D). These might be used because they support a wider range of plane wave angles for an enclosed control region than a finite plane or line array. In listening spaces it is usually not practical to locate sources beneath the listeners, so dome shaped boundaries are common. DMC can be used to focus the control where it is needed most. In an auditorium this might be a bar shaped region raised off the ground, enclosing the heads and shoulders of the listeners, as shown in Fig. 17. Fig 18 shows a sideways directed plane wave. Source energy is low outside the beam, so this reduces unnecessary excitation of the enclosing space. The previous discussion on distribution of sources is relevant 3.2.4, although optimal redistribution is not attempted here.

3.2.6 Separated regions

When control regions are separated according to the $N = kr$ criterion, a different target field can be applied to each. Following the previous result for a single region, the least number of sources required is expected to be approximately $J = \sum 2P/\lambda$ or $\sum 4A/\lambda^2$, where P, A are the perimeters or sur-

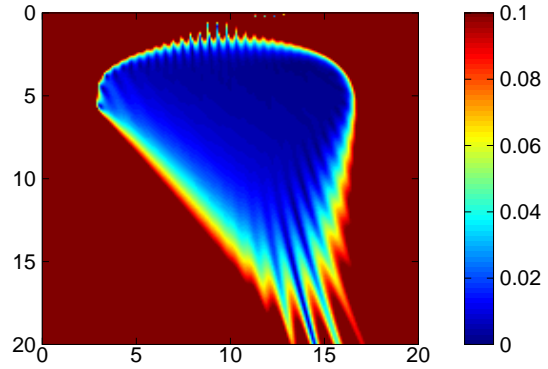
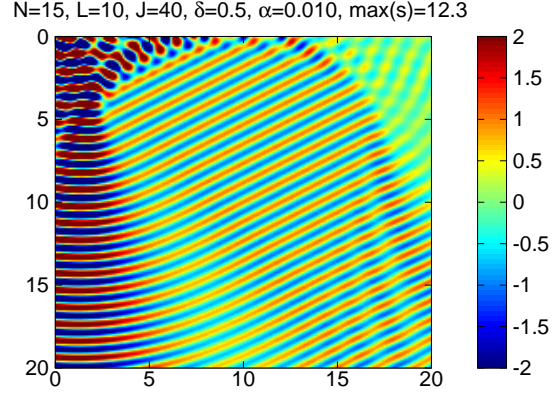


Figure 16: Reconstruction of a plane wave from a line array above.

face areas of control regions of any shape. This is found to be true for regions that are constrained to independent fields. Fig. 19 shows an example for 4 constraint regions with $N = 7, 9, 10, 8$ and $J = 100$. The third region is a quiet zone, and the error shown is absolute, not relative as shown for the other regions. The expected number of required sources is $\sum 2P/\lambda = \sum 2N_i = 68$. This is found to be sufficient, and for $J < 68$ the solution progressively degrades.

Where some of the subregions sample the same well behaved underlying field, then the number of sources required can be less than that given by the surface sum. This is because two such subregions

N=3, L=10, J=30, $\delta=0.5$, $\alpha=0.100$, $\max(s)=1.9$

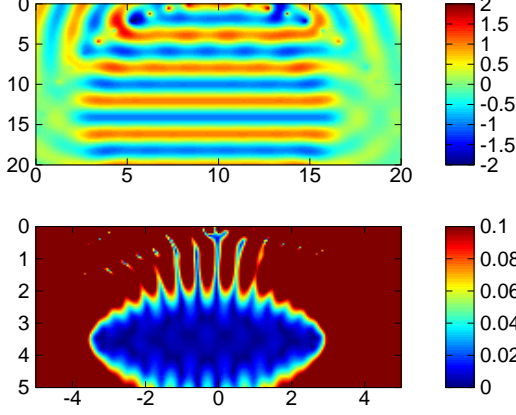


Figure 17: Semicircular boundary, wave from above.

N=3, L=10, J=30, $\delta=0.5$, $\alpha=0.100$, $\max(s)=2.2$

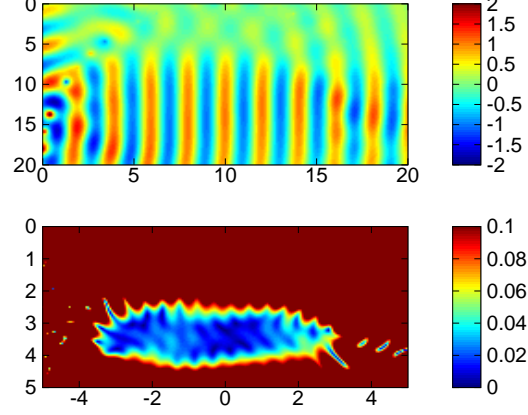


Figure 18: Semicircular boundary, wave from the side.

may be a subset of a region with a surface that is smaller than the surface sum of the members.

Some further tests were made to investigate how the relative position of regions and their targets affect the solution. The results are summarized here. First two single centre regions were considered with $N = 10$, spaced 4λ apart inside a 5λ by 10λ boundary. Therefore each region has a radius according to $N = kr$ of $10/(2\pi)\lambda \approx \lambda$. With each target set to a plane wave in the same direction between centres, low error is achieved for 44 sources, $\max(p) = 1.4$, $\alpha = 0.01$. With target waves pointing inwards against each other, low error is achieved for 44 sources, $\max(p) = 145$, $\alpha = 0.0001$. Increasing the number of sources to 300 brings $\max(p)$ down to a limit of ≈ 52 . The lower $\max(p)$ limit is caused by conflicting targets, rather than lack of sources.

The interior was then expanded to 10λ by 10λ with the target waves pointing in opposite directions sideways. For 40 sources a low error solution was found for $\max(p) = 3$. However if the regions are extended sideways into bars, they become difficult to control. Bars of width 6λ do not have good solutions for any set of parameters. In Fig. 19 the modal regions are close, but none of the regions lies in the path of reproduced plane waves of another region, and so a much better solution is achieved than would otherwise be possible.

In summary, in some cases there are no low error

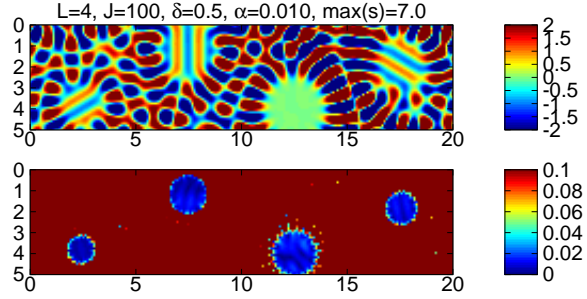


Figure 19: Reconstruction of four independent regions in a rectangular array.

solutions for any number of sources. Even if a low error solution exists then a low error, low $\max(p)$ solution may not exist. Low error, low $\max(p)$ solutions generally require at most $J = \sum 2P/\lambda$ or $\sum 4A/\lambda^2$ sources.

In [26] the possibility of moving a single modal centre in realtime to track the head of a listener is examined. With multiple modal centres multiple listeners can be tracked. If the listeners are all listening to the same underlying field then this can be achieved with a $\delta = 0.5$ boundary. Sources cannot be repositioned in real-time, so a source reduction is not possible while also allowing free movement of the listeners. If the listeners each have a different target then control in general is impractical unless the listeners remain well separated.

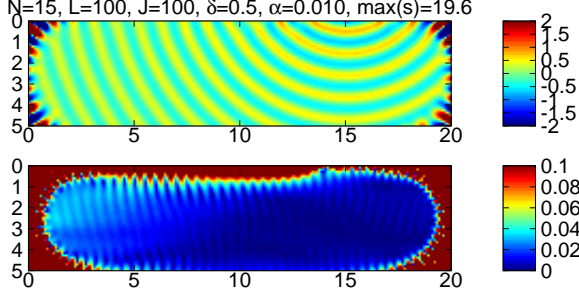


Figure 20: Reconstruction of an exterior point source.

3.2.7 Point source control methods

The reproduction of free field regions of point source fields is now considered. The modal region expansions for a simple source are given in Section 3.1. An example is shown in Fig. 20 for a source $4H^{(2)}(r)$ located at $(15\lambda, -2.5\lambda)$. It was noted previously that a sample field becomes progressively more badly behaved as it approaches a source. This is evident in the example where a reduction of δ to 0.4 causes the error and $\max(p)$ to fall to low values.

For an internal source it is not possible to include the centre in a constrained region because the area would then not be a free-field. However constrained regions can be placed around the centre. It is then possible to have several listeners facing each other while perceiving a single source at the centre as they would with a real central source field, Fig. 21. By comparison, in WFS or HOA listeners must be on one side. The target here is $10H^{(2)}(r)$, and has been amplified to give a clearer picture for the scale used in the Figures.

This is a special case of independent control regions sampling an underlying field as discussed in the last section. It is perhaps surprising that energy is able to flow inwards from multiple directions then outwards through the control regions, requiring only moderate driver and field strength, except at localised field strength hot spots apparent in Fig. 22, which shows the absolute value of the field.

If the control regions are enlarged slightly, the solution rapidly disintegrates, even with reduced α and δ . A geometrical limit is reached.

The question arises how far can the target source

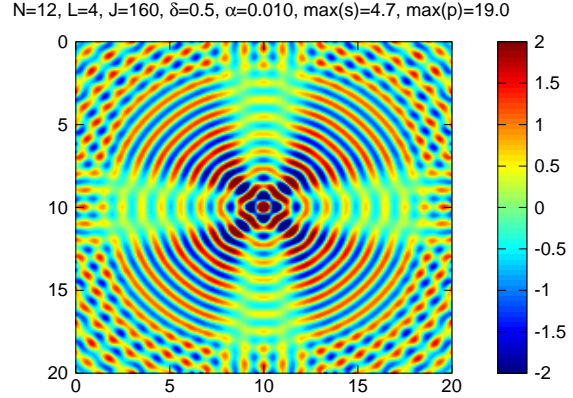


Figure 21: Interior target source surrounded by separate constrained regions.

be enclosed in a single region, maybe more than a half-space? This would be counterintuitive as this limit is a half-space for an FBE of a source or a focused source. A constrained arc of overlapping modal regions was constructed around a source such that the arc has an angle $> \pi$ to the centre. The fountain-like solution shown in Fig. 21 is achieved with low α and high $\max(p)$. The error scale maximum has been changed to 0.5. The monopole is clearly over an arc $> \pi$, but this is not a practical solution. Attempts to extend the arc further lead to increased error, regardless of parameter choices.

$N=12, L=4, J=160, \delta=0.5, \alpha=0.010, \max(s)=4.7, \max(p)=19.0$

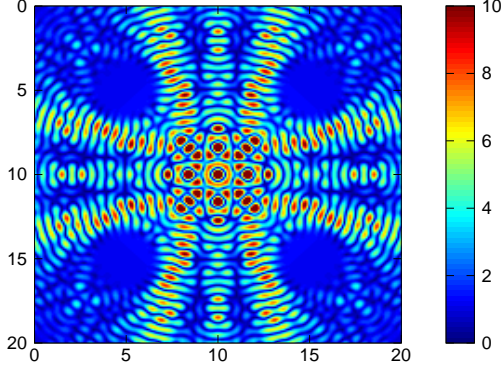


Figure 22: Absolute value plot of an interior target source surrounded by separate regions.

3.3 Encoding and decoding

Driver signals provide a direct encoding for the reproduced sound field. If the least number of optimally placed drivers are used to achieve the desired accuracy for the range of possible target fields, then the driver signals provide an optimally compact encoding. If the boundary is concave it cannot be replaced by an enclosing convex boundary with a fewer drivers, because there exist fields that cannot be reproduced by the convex boundary that can be reproduced by the concave boundary. An example is the field of an external source placed near the concave point, Fig. 24. The study connected with Fig. 23 shows that it is not possible to reproduce such a field as part of a free field.

Suppose that given the driving functions $s_j^{(1)}$ for one array we wish to find the driving functions for another $s_j^{(2)}$ that reproduces all or part of the first. First the control regions must be oriented so that an intersection region can be defined, as shown in Fig. 25. The second array can reproduce anything in the overlap that can be reproduced by the first. The j th driver in the first array defines a field in the intersection with a reproduction in the second array that is defined here as a filter matrix s_{kj} . Then for a general field given by driving functions $s_j^{(1)}$, the sources reproducing this in the intersection on the second array are $s_k^{(2)} = \sum s_{kj} s_j^{(1)}$.

Ambisonics uses an array independent encoding, the FBE. A single FBE can also be used for a general convex control region by finding a minimal en-

$N=8, L=8, J=160, \delta=0.5, \alpha=0.001, \max(s)=127.2, \max(p)=411.6$

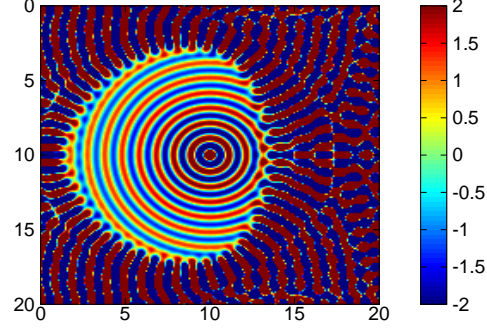
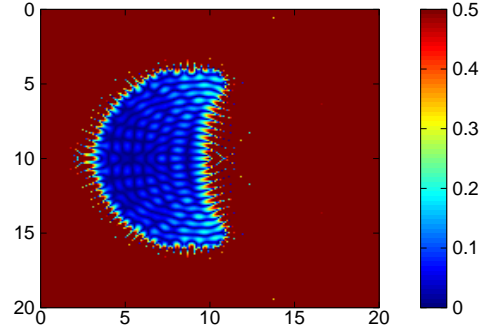


Figure 23: Interior target source surrounded by a control arc with angle to centre $> \pi$.



closing FBE region, Fig. 26. This in general will have redundancy because it encodes some regions that are not reproduced.

FBE encoding is not in general possible for a concave boundary. An external source at the concave point cannot be represented as part of a free field. Other encoding methods such as driver encoding are necessary.

For each FBE basis field, there are driving functions reproducing the field in the control region, found using DMC. Each basis has an expansion at each modal centre given by its translation from the encoding FBE. The translations can be given in closed form, as for example in [27]. Call these the decoding coefficients, B_{nj} for the n th harmonic and j th driver. Then the driving functions for a general FBE with coefficients b_n are $s_j = \sum b_n B_{nj}$.

HOA supports every direction of plane wave ex-

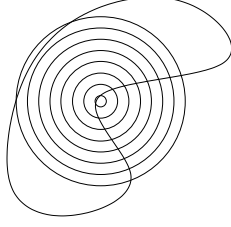


Figure 24: Source at concave corner of boundary.

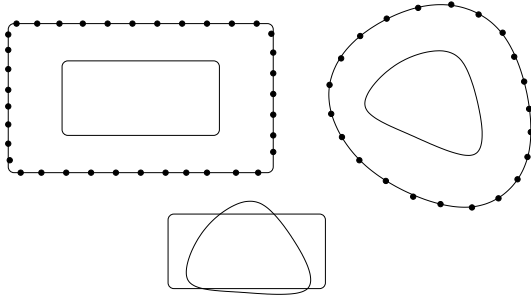


Figure 25: Two loudspeaker arrays with control regions, and the control region intersection.

actly up to the angular bandlimit. Open boundaries can support a restricted set of plane waves accurately over the target region, although generally not exactly in the bandlimited sense. Examples have been given using DMC. WFS on a closed boundary is effectively of this type, since drivers that face against the plane wave have no signal.

The problem of decoding for restricted arrays was considered previously [28]. The approach described here is based on a plane wave decomposition. How should a general HOA encoding be decoded to a restricted array? The simplest strategy is to map plane waves that fall outside the set to silence. More complex strategies are conceivable in which restricted directions are mapped to permitted directions, or mixes of directions. For the first strategy a restricted set of basis fields $\bar{R}_n(\mathbf{r})$ can be constructed by plane wave re-expansion of the HOA basis, with angular windowing [29]. The angular sampling resolution must be sufficient to represent plane waves accurately over the whole reproduction region. The restricted decoding functions \bar{B}_{nj} , with j indexing the drivers, are then found by summing the decoding contribution from each plane wave, for example using DMC.

If a target field with HOA coefficients b_n has no

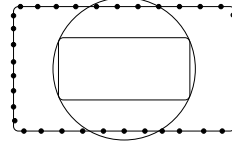


Figure 26: Single FBE region enclosing the control region.

energy in the restricted directions then

$$\sum b_n R_n(\mathbf{r}) = \sum b_n \bar{R}_n(\mathbf{r}) \quad , \quad (11)$$

even though R_n are unrestricted. This means that a restricted target field will be accurately reproduced on the restricted array when driven by sources $s_j = \sum b_n \bar{B}_{nj}$. Other field encodings will lose restricted plane wave components when reproduced on the restricted array.

In [28] a reduced size basis is constructed in order to reduce the size of encodings. An alternative approach is to work directly with the plane wave decomposition, discarding plane wave directions that cannot be reproduced. The method can be adapted to encode efficiently for an arbitrary reproduction region by varying the angular sample spacing as $\lambda/2r$, where r is the distance from the origin to the reproduction boundary. This is related to encoding using optimal driving signals, as discussed earlier.

4 Conclusion

DMC has been shown to be effective for a variety of harmonic target fields. In particular the two initial motivating cases, general sub-regions and concave boundaries have been demonstrated. The investigation of interior point sources showed that constraint regions can be used as a design method for the global field, rather than just a control method, sometimes producing unexpected and useful results. This could be used to generate encodings that can be applied across different reproduction systems.

The optimal source spacing for total interior control was found to be near $\lambda/2$ in every case. The spacing must be reduced around sharp corners, particularly concave points, and for fields that are not well behaved in the sense given. Interior regions sampling the same well behaved underlying field are

controllable using at most $2P/\lambda$ or $4A/\lambda^2$ sources in 2D or 3D respectively for perimeter P and surface area A . Independent subregions are not in general practically controllable, but are when sufficiently separated.

DMC driving functions are straightforward to evaluate, but not directly at interactive rates. If driving functions for HOA or plane wave basis functions were pre-calculated, then the driving functions for sources could be calculated at interactive rates. The application of DMC driving functions as filters is costly compared with those required by WFS, and HOA on a circle or sphere, but feasible in real-time. Playback of pre-calculated driving signals from hard-disc storage is possible, and does not critically depend on driver calculation costs. Similarly HOA encodings can be pre-calculated, and decoded with interactive rotation.

Areas for future investigation include time domain and wideband treatment including aliasing. The freedom to specify many constraints may allow aliasing to be optimized in some sense. The more practical case of 3D sources for horizontal arrays should be considered, as well as a more detailed study of optimal driver placement. Driver placement may also be relevant for alias control. Complex drivers were mentioned. These can be studied from the viewpoint of realistic driver modelling and also the control of external radiation. The method of regularization can be generalized so that varying priority can be given to different constraints. An interactive graphical environment for constraint design could be a useful tool.

References

- [1] D. Menzies, "Sound synthesis for general enclosures," in *Proc. 2nd Int. Symposium on Ambisonics and Spherical Acoustics, IRCAM Paris*, May 2010.
- [2] M. A. Gerzon, "Periphony (with-height sound reproduction)," in *Proc. AES 2nd Convention, Munich*, 1972.
- [3] J. Daniel, "Spatial sound encoding including near field effect," in *Proc. AES 23rd International Conference, Helsinger, Denmark*, 2003.
- [4] M. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Soc.*, vol. 53, no. 11, pp. 1004–1025, November 2005.
- [5] E. Williams, *Fourier Acoustics: sound radiation and nearfield acoustical holography*. Elsevier, 1999.
- [6] D. Menzies and M. Al-Akaidi, "Ambisonic synthesis of complex sources," *Journal of the Audio Engineering Society*, vol. 55, no. 10, pp. 864–876, October 2007.
- [7] S. Spors and R. Rabenstein, "Spatial aliasing artifacts produced by linear and circular loudspeaker arrays spatial aliasing artifacts used for wave field synthesis," in *Preprint 6671, AES 120th Convention, Paris*, May 2006.
- [8] M. Noisternig, A. Sontacchi, T. Musil, and R. Höldrich, "A 3d ambisonic based binaural sound reproduction system," in *In Proc. AES 24th International Conference: Multichannel Audio, The New Reality, Banff*, 2003.
- [9] M. A. Poletti and T. D. Abhayapala, "Interior and exterior sound field control using general two-dimensional first-order sources," *J. Acoust. Soc. Am.*, vol. 129, no. 1, pp. 234–44, 2011.
- [10] M. Poletti, "An investigation of 2d multizone surround," in *Proc. AES 125th Convention, San Francisco*, 2008.
- [11] Y. Wu and T. Abhayapala, "Spatial multizone soundfield reproduction," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei*, April 2009, pp. 93–96.
- [12] A. J. Berkhout and P. de Vries, D. and Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, pp. 2764–2778, 1993.
- [13] K. Brandenburg, S. Brix, and T. Sporer, "Wavefield synthesis: from research to applications," in *In Proc. EUPISCO, Vienna*, 2004.
- [14] E. Verheijen, "Sound reproduction by wave field synthesis," Ph.D. dissertation, Delft University of Technology, January 1998.

- [15] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *Preprint 7358, AES 124th Convention, Amsterdam*, May 2008.
- [16] J. Ahrens and S. Spors, "Spatial encoding and decoding of focused virtual sound sources," in *Proc. Ambisonics Symposium, Graz*, 2009.
- [17] D. Menzies, "Calculation of near-field head related transfer functions using point source representations," in *Proc. Ambisonics Symposium, Graz*, 2009.
- [18] S. Spors and J. Ahrens, "Spatial sampling artifacts of focused sources in wave field synthesis," in *In Proc. Intern. Conf. on Acoustics NAG/DAGA Rotterdam*, March 2009.
- [19] —, "Local sound field synthesis by virtual secondary sources," in *In Proc. AES 40th International Conference, Tokyo*, 2010.
- [20] F. Fazi, "Surround system based on three dimensional sound field reconstruction," in *Proc. AES 125th Convention, San Francisco*, 2008.
- [21] Y. Liu, *Fast Multipole Boundary Element Method - Theory and Applications in Engineering*. Cambridge University Press, 2009.
- [22] O. Kirkeby and P. A. Nelson, "Reproduction of plane wave sound fields," *Journal of the Acoustical Society of America*, vol. 94, no. 5, pp. 2992–3000, 1993.
- [23] P.-A. Gauthier and A. Berry, "Adaptive wave field synthesis with independent radiation mode control for active sound field reproduction," *Journal of the Acoustical Society of America*, vol. 119, no. 5, pp. 2721–2737, 2006.
- [24] S. Spors, H. Buchner, and R. Rabenstein, "A novel approach to active listening room compensation for wave field synthesis using wave-domain adaptive filtering," in *In Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Montreal*, 2004.
- [25] P. Morse and K. Ingard, *Theoretical Acoustics*. McGraw-Hill, 1968.
- [26] J. Hannemann and K. Donohue, "Virtual sound source rendering using a multipole-expansion and method-of-moments approach," *J. Audio Eng. Soc.*, vol. 56, no. 6, pp. 473–481, June 2008.
- [27] N. A. Gumerov and R. Duraiswami, *Fast multipole methods for the Helmholtz equation in three dimensions*. Elsevier Science, 2005.
- [28] H. Pomberger and F. Zotter, "An ambisonics format for flexible playback layouts," in *Proc. Ambisonics Symposium, Graz*, 2009.
- [29] D. Menzies and M. Al-Akaidi, "Nearfield binaural synthesis and ambisonics," *Journal of the Acoustical Society of America*, vol. 121, no. 3, pp. 1559–1563, 2007.