

Planarity panning for listener-centered spatial audio

Philip Coleman¹, Philip J. B. Jackson¹

¹*Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, Surrey, GU2 7XH, UK*

Correspondence should be addressed to Philip Coleman (p.d.coleman@surrey.ac.uk)

ABSTRACT

Techniques such as multi-point optimization, wave field synthesis and ambisonics attempt to create spatial effects by synthesizing a sound field over a listening region. In this paper, we propose planarity panning, which uses superdirective microphone array beamforming to focus the sound from the specified direction, as an alternative approach. Simulations compare performance against existing strategies, considering the cases where the listener is central and non-central in relation to a 60 channel circular loudspeaker array. Planarity panning requires low control effort and provides high sound field planarity over a large frequency range, when the zone positions match the target regions specified for the filter calculations. Future work should implement and validate the perceptual properties of the method.

1. INTRODUCTION

Approaches to spatial sound reproduction have been investigated over many years. For instance, sound field synthesis (SFS) approaches aim to physically recreate sound fields due to virtual sources. Such approaches include Higher-order ambisonics (HOA) [1, 2], Wave field synthesis (WFS) [3, 4], and least-squares pressure-matching (PM) techniques [5, 6]. These methods were primarily developed in order to advance spatial audio reproduction from stereophony towards the situation where any auditory scene could be created for a listener. Alternatively, amplitude panning (AP) may be used to place a source, using up to three loudspeakers [7]. A thorough overview of the development of spatial audio technologies through to the present day is presented in [8].

Factors affecting the selection of reproduction approach include the number of loudspeakers available, the desired area over which control is effected, and the simplicity of the signal processing required to achieve the desired localization. For instance, AP and ‘traditional’ 0th or 1st order ambisonics use relatively few loudspeakers, have a very small ‘sweet spot’ for the optimal spatial impression, and the processing is relatively straightforward. On the other hand, WFS and HOA aim to extend the reproduction over a wide region, but require many loudspeakers for broadband reproduction, and the necessary processing is increased. In addition to attempting the syn-

thesis of a sound field inside an entire array, many investigations have focused on reproduction over a localized region, in which the listener is to be located. Such an approach is useful when the listening position is fairly well known (or can be tracked), and the listening zone may be designed so that it is large enough to comfortably contain the listener’s head. In these scenarios, where relatively few virtual or physical microphones are required to sample the sound field with a suitable density, multi-point optimization methods such as PM represent a compelling choice. These methods are based on simulated or measured transfer responses between the loudspeakers and microphones, which can be straightforwardly obtained, and also carry room information, meaning that room reflections may be directly accounted for in the optimization (the listening room can also be compensated in the analytical approaches [9, 10, 11]).

One well-known restriction on SFS systems is the effect of spatial aliasing, the frequency of which depends on the loudspeaker density. In particular, WFS and HOA suffer from aliasing effects at high frequencies, and each method has characteristic artifacts which have been previously compared for spatial audio [8, 12]. For HOA, the aliasing occurs when there is an insufficient order of basis function expansion to cover the desired area, leading to a reduction in the size of the region of accurate reproduction, although if the listener is in the sweet spot

they may not notice these effects until a much higher frequency. On the other hand aliasing occurs for WFS when the loudspeaker spacing is too wide to reproduce the desired frequency, and strong grating lobes appear, affecting the sweet spot in addition to the rest of the sound field. The multi-point approaches are subject to the same physical constraints, with properties typically between HOA and WFS [8].

Multi-point optimization has recently been applied to investigations of personal audio [13]. In this context, the planarity control cost function was introduced in [14], creating a high level difference between the listening zones and using a spatial constraint to reduce self-cancellation artifacts typically present in energy cancellation systems. Here, we present simulation results whereby the spatial constraint is formulated for single-zone spatial audio reproduction, referring to the technique as planarity panning (PP). The method is rather different from typical SFS approaches in that the sound propagation across the listening region is not explicitly controlled or optimized. Rather, similar to [15], the spatial effect is achieved by projecting the sound pressures in to a spatial domain (using microphone array beamforming), and narrowing the range of possible directions from which the energy may impinge on the zone. Although 2-D simulations are presented here, the concept straightforwardly extends to 3-D. Unlike WFS and HOA, there is no restriction on the loudspeaker positions, and unlike HOA, there is no restriction on the microphone array geometry for measured implementations. Furthermore, the method is highly efficient, increasing robustness to room reflections in a practical system.

In this paper, the background and theory of some existing spatial reproduction techniques are first introduced. Then, PP is described. Simulations are presented to demonstrate the performance of PP for reproducing 5 virtual sources (in locations corresponding to 5.0 surround sound channels), at central and off-central listening positions, and the method is compared with PM, AP, WFS and HOA. Finally, conclusions are drawn.

2. BACKGROUND

In the following, our notation is introduced, and existing spatial sound approaches are briefly described.

2.1. Notation and fundamentals

SFS requires finding a solution to the interior reproduction problem, illustrated in Fig. 1. The sound field is

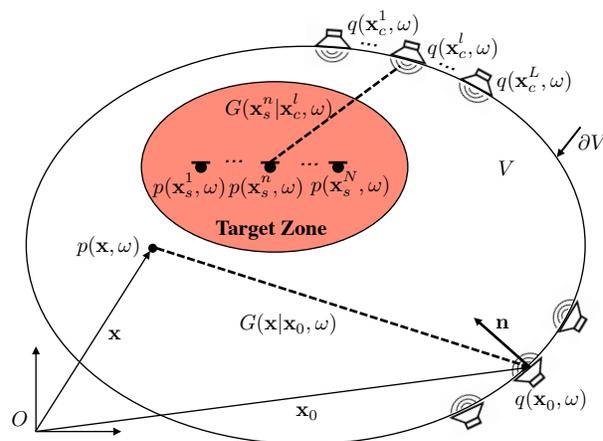


Fig. 1: Illustration of the interior reproduction problem and the geometry of a multi-point system.

reproduced over the source-free volume V . The pressure at a certain observation point \mathbf{x} at angular frequency $\omega = 2\pi f$ is indicated by $p(\mathbf{x}, \omega)$. The position of a certain point on the surface ∂V is defined as \mathbf{x}_0 , and the inward pointing surface normal at \mathbf{x}_0 is indicated by \mathbf{n} . A source at \mathbf{x}_0 has source weight $q(\mathbf{x}_0, \omega)$. The Kirchhoff-Helmholtz integral represents solutions of the Helmholtz equation with inhomogenous boundary conditions and states that the sound field at any point $\mathbf{x} \in V$ is uniquely determined by the sound pressure and inward facing sound pressure gradient on the boundary ∂V [16]. In practice, two modifications must be made in order to derive the loudspeaker weights. Firstly, the Kirchhoff-Helmholtz integral defines the whole sound field including the infinite region outside of the volume of interest. This means that control of either the pressure or pressure gradient around ∂V is adequate to reproduce the sound field in V [16]. Usually, a single layer potential of monopoles is used, as these are simpler and represent real loudspeakers relatively well [12]. The sound pressure produced within V by the continuous layer of monopoles can be written in terms of the source weights of the monopoles as [12]

$$p(\mathbf{x}, \omega) = - \oint_{\partial V} G(\mathbf{x}|\mathbf{x}_0, \omega) q(\mathbf{x}_0, \omega) dA(\mathbf{x}_0), \quad (1)$$

where $G(\mathbf{x}|\mathbf{x}_0)$ represents the free-field Green's function, $dA(\mathbf{x}_0)$ is an infinitesimal surface element of ∂V , and the problem is to select $q(\mathbf{x}_0, \omega)$ for each position \mathbf{x}_0 . For the second modification, the assumption of a continuous

layer of monopole sources must be violated. WFS and HOA represent different approaches to making these two modifications and will be briefly introduced below.

Figure 1 also shows the notation used for a multi-point approach to reproduction, where the subscripts c and s on the position vectors denote control sources and sensors, respectively. The superscripts identify the individual microphones and loudspeakers, which may be placed arbitrarily in the room, and the system is defined acoustically. However, the synthesized sound field is not spatially continuous. For our anechoic simulations, the free-field Green's function is used, but measured room impulse responses are easily adopted. For each frequency, the source weights can be written in vector notation as $\mathbf{q} = [q(\mathbf{x}_c^1), \dots, q(\mathbf{x}_c^L)]^T$, where there are L loudspeakers and $q(\mathbf{x}_c^l)$ is the complex source weight of the l th loudspeaker, positioned at \mathbf{x}_c^l . Similarly, the complex pressures at the microphones are written as $\mathbf{p} = [p(\mathbf{x}_s^1), \dots, p(\mathbf{x}_s^N)]^T$, where there are N control microphones and $p(\mathbf{x}_s^n)$ is the complex pressure at the n th microphone, positioned at \mathbf{x}_s^n . The plant matrix contains the transfer functions between each loudspeaker and microphone, and is defined as:

$$\mathbf{G} = \begin{pmatrix} G(\mathbf{x}_s^1|\mathbf{x}_c^1) & \cdots & G(\mathbf{x}_s^1|\mathbf{x}_c^L) \\ \vdots & \ddots & \vdots \\ G(\mathbf{x}_s^N|\mathbf{x}_c^1) & \cdots & G(\mathbf{x}_s^N|\mathbf{x}_c^L) \end{pmatrix}, \quad (2)$$

where $G(\mathbf{x}_s^n|\mathbf{x}_c^l)$ is the transfer function between the n th microphone and the l th loudspeaker. The pressures at the microphone positions may be written as $\mathbf{p} = \mathbf{G}\mathbf{q}$, which is a discretized version of Eq. (1) for all microphone positions.

2.2. Higher-order ambisonics

Derivation of the source weights by HOA depends on the explicit solution of Eq. (1), which is a compact Fredholm operator of zero index. A solution is given by expanding each element of Eq. (1) in to a series of orthogonal basis functions [8]. The source weights are expressed as [2]

$$q(\mathbf{x}_0, \omega) = \sum_{n=1}^{N_m} \tilde{q}_n(\omega) \psi_n(\mathbf{x}), \quad (3)$$

where $\psi_n(\mathbf{x})$ are the basis functions, N_m is the order of the expansion, and $\tilde{q}_n(\omega)$ is the projection of the source weights on to the basis functions. The size of sweet spot increases with the order of expansion, which is determined by the available number of loudspeakers.

2.3. Wave field synthesis

The WFS approach is usually defined in terms of Rayleigh's first integral [16], which states that the sound pressure in one half-space (the 'target' half-space) can be specified by a continuous distribution of monopole sources along an infinite planar boundary. It is generally assumed that a bent surface can be approximated as a series of planar ones [8, 12]. One result of this assumption is that in WFS, sources whose normal \mathbf{n} has a negative component in the propagation direction of the desired wave field are often switched off [12]. Accordingly, a window function $w(\mathbf{x}_0)$ is introduced, and the modified Rayleigh integral is [8]:

$$p(\mathbf{x}, \omega) = - \oint_{\partial V} 2w(\mathbf{x}_0) \frac{\partial}{\partial \mathbf{n}} p(\mathbf{x}_0, \omega) G(\mathbf{x}|\mathbf{x}_0, \omega) dA(\mathbf{x}_0). \quad (4)$$

The source weight can be simply derived by comparing Eqs. (1) and (4):

$$q(\mathbf{x}_0, \omega) = -2w(\mathbf{x}_0) \frac{\partial}{\partial \mathbf{n}} p(\mathbf{x}_0, \omega). \quad (5)$$

Complex wave-fields may be reproduced by WFS by prior decomposition in to plane-wave components which can then be reproduced by subsets of the loudspeakers [17]. One consequence of the unwrapping of the planar boundary around an arbitrary shape ∂V is that exact sound field reproduction is not possible within V using WFS. Implementations of WFS therefore tend to be constrained to circular, planar or square array geometries [12]. However, a number of experimental and commercial WFS systems have been successfully realized [8]. As the basis of WFS depends on the plane-wave representation at the boundary, the aliasing artifacts are strongly linked to the loudspeaker spacing, and affect the whole of V (compared to the shrinking of the sweet spot in HOA).

2.4. Least-squares pressure matching

Virtual sources can also be synthesized by least-squares optimization. For plane wave sound fields, the desired field $\mathbf{d} = D e^{j k \mathbf{x}_s^n \cdot \mathbf{u}_\phi}$, for $n = 1, 2, \dots, N$, where D gives the pressure amplitude, \mathbf{x}_s^n is the position of the n th control microphone, \cdot denotes the inner product, and \mathbf{u}_ϕ is the unit vector in the direction of the incoming plane wave. The cost function, with a constraint to fix the effort to a certain Q , is [6]:

$$J = (\mathbf{p} - \mathbf{d})^H (\mathbf{p} - \mathbf{d}) + \lambda (\mathbf{q}^H \mathbf{q} - Q). \quad (6)$$

Using Lagrange multipliers, the solution can be found by taking the derivatives with respect to \mathbf{q} and λ :

$$\mathbf{q} = (\mathbf{G}^H \mathbf{G} + \lambda \mathbf{I})^{-1} \mathbf{G}^H \mathbf{d}; \quad \mathbf{q}^H \mathbf{q} = Q. \quad (7)$$

The Lagrange multiplier λ is numerically chosen to satisfy the control effort constraint, which is here set to correspond to 0 dB (cf. Eq. (12)).

2.5. Amplitude Panning

Finally, the rather simple approach of amplitude panning can be used. In 2-D, a virtual source is placed by adjusting the gains of a pair of loudspeakers surrounding the intended virtual source location φ by the panning law [7]:

$$\frac{\tan \varphi_s}{\tan \varphi_0} = \frac{g_1 - g_2}{g_1 + g_2}, \quad (8)$$

where φ_s is the virtual source angle relative to the central axis of the loudspeaker pair, and $2\varphi_0$ is the angle between the pair of loudspeakers.

3. PLANARITY PANNING

Rather than synthesizing the exact sound field over the reproduction area, planarity panning uses a zone-based approach that, as with the least-squares approach, depends on an array of microphones sampling the target region. The energy in the target region is projected in to a spatial domain by superdirective beamforming, and in this domain the energy in the zone is optimized. Similar work in [15] reproduced a plane wave by energy focusing, exploiting the appearance of a plane wave as a point in the wavenumber domain. The approach was found to improve precision of plane wave placement with respect to HOA and require fewer loudspeakers than WFS [18].

The planarity panning cost function is given as a constrained maximization, similar to [19]:

$$J = \mathbf{p}^H \mathbf{Y}^H \mathbf{\Gamma} \mathbf{Y} \mathbf{p} - \lambda (\mathbf{q}^H \mathbf{q} - Q), \quad (9)$$

where λ is a Lagrange multiplier, the $I \times N$ steering matrix \mathbf{Y} projects the sound pressure at the control microphones in to a spatial domain, and Q is a constraint on the array effort. Equation (9) can be interpreted as the maximization of acoustic brightness, as [19] but via the spatial domain, and constrained by a certain sum of squared source weights. The diagonal matrix $\mathbf{\Gamma}$ allows a weighting to be applied based on the desired incoming plane wave directions:

$$\mathbf{\Gamma} = \text{diag}[\gamma_1, \gamma_2, \dots, \gamma_I], \quad (10)$$

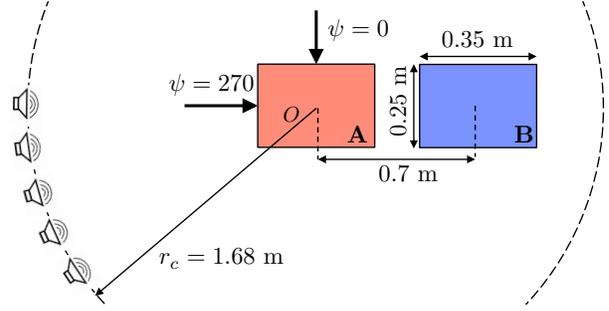


Fig. 2: Simulation geometry showing the central portion of a circular array, with central (A) and side (B) listening positions. The angles of energy impinging on to the zone are also shown (shifted for side position).

where $0 \leq \gamma_i \leq 1$ is the weighting corresponding to the i th steering angle. Energy will therefore be focused in the direction of the nonzero elements of $\mathbf{\Gamma}$. The steering matrix \mathbf{Y} is populated by a regularized max-SNR beamformer with fixed beam-width [14, 13, 20], mapping between the observed microphone pressures and the energy components, $\mathbf{w} = [w_1, \dots, w_I]^T = |\mathbf{Y}\mathbf{p}|^2$. The diagonal elements γ_i are set with $\gamma_\varphi = 1$ and a raised-cosine window of 5 degrees for $i = \varphi \pm 5$ [14]. For comparison, the method of [15] can be expressed by Eq. (9), with \mathbf{Y} populated by spatial Fourier transform and $\mathbf{\Gamma}$ having infinitesimal angular resolution with a single nonzero element γ_φ at the desired plane wave direction.

The point that maximizes J can be found by setting its derivatives with respect to \mathbf{q} and λ , respectively, to zero:

$$\mathbf{G}^H \mathbf{Y}^H \mathbf{\Gamma} \mathbf{Y} \mathbf{G} \mathbf{q} = \lambda \mathbf{q}; \quad \mathbf{q}^H \mathbf{q} = Q. \quad (11)$$

The derivative $\partial J / \partial \mathbf{q}$ describes an eigenvalue problem, and the optimal source weight vector \mathbf{q} is proportional to the eigenvector $\hat{\mathbf{q}}$ corresponding to the maximum eigenvalue of $\mathbf{G}^H \mathbf{Y}^H \mathbf{\Gamma} \mathbf{Y} \mathbf{G}$. The derivative $\partial J / \partial \lambda$ is used to enforce the effort constraint Q (corresponding to $E = 0$ dB, Eq. (12)). Thus, PP maximizes the sound pressure level (SPL) into the target direction for a certain input power. By scaling the source weights, one can set either the effort or the brightness (i.e., the target SPL in the bright zone).

4. SIMULATIONS

Computer simulations were conducted to establish the performance of PP, and provide a comparison with the

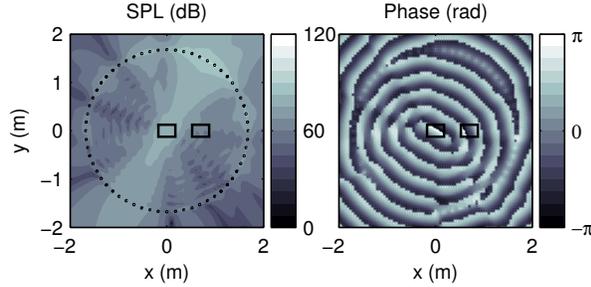


Fig. 3: Phase and sound pressure level in the simulated room at 1 kHz, using PP ($L = 60$) to create a virtual loudspeaker at 30 degrees for a listener in the centre zone (A).

methods described in Section 2. A 60 channel circular array of radius 1.68 m was used, and two $0.25 \text{ m} \times 0.35 \text{ m}$ listening positions were evaluated, one at the centre of the circle (A), and the other offset by 0.7 m (B). The system geometry is shown in Fig. 2. For the multi-point approaches, 192 virtual microphones were used in each listening region, arranged as a 2.5 cm spaced grid. For experiments considering the effect of the number of loudspeakers on performance, regularly-spaced subsets of 15, 20 and 30 loudspeakers were used. Virtual far field sources (plane-wave fields) at the 5.0 surround sound positions (i.e. 0, -30, 30, -110, 110 degrees) were reproduced using each method. Simulations were conducted at frequencies between 250–6000 Hz, at 250 Hz intervals. The array aliasing limits, based on the spacing of the loudspeakers around the circular array, were 250–975 Hz, for 15–60 loudspeakers, and the upper order of harmonic expansion for HOA was $\lfloor L/2 \rfloor$. The source weights for WFS and HOA were calculated using the Sound Field Synthesis Toolbox [21], and evaluated by applying the weights in our own software environment. For off-centre reproduction, the coordinates for the array centre and amplitude normalization point were adjusted. The evaluation metrics of control effort and planarity were used [13], in addition to the RMSE of the direction of the principal energy component impinging on the listening region. The control effort E is the energy that the loudspeaker array requires, relative to a reference source q_r producing the same pressure in the listening zone, and planarity η is the extent to which the sound field in the listening zone resembles a plane wave [20]. The metrics are defined as:

$$E = 10 \log_{10} \left(\frac{\mathbf{q}^H \mathbf{q}}{|q_r|^2} \right); \quad \eta = \frac{\sum_i w_i \mathbf{u}_i \cdot \mathbf{u}_\alpha}{\sum_i w_i}, \quad (12)$$

where \mathbf{u}_i is the unit vector associated with the i th component's direction, \mathbf{u}_α is the unit vector in the principal direction $\alpha = \arg \max_i w_i$, and \cdot denotes the inner product. The beamforming approach was also used to assess the direction of arrival of the sound impinging on the listening zone in terms of the root-mean-square error (RMSE):

$$\varepsilon = \sqrt{\frac{1}{F} \sum_{f=1}^F |\alpha(f) - \varphi|^2}, \quad (13)$$

where F is the number of frequencies considered.

4.1. Properties of planarity panning

Sound fields at 1 kHz using the full 60 channel array to render the virtual channel at 30 degrees in position A are shown in Fig. 3, showing the phase and SPL distributions in the simulated room. It can be seen from the phase plots that a planar sound field is reproduced, with the energy impinging from the intended direction. The results generalize well across virtual channel position and frequency, with the minimum RMSE of 0.6 degrees over 250–6000 Hz for the frontal virtual source position, and the highest RMSE of 2.5 degrees for the left- and right-surround positions over the same frequency range. These values are within the tolerance of the standard 5.0 surround positions. When the number of loudspeakers are reduced, the RMSEs increase, although the scores generally remain below 5 degrees (15 loudspeaker surround channels 8.4 degrees), and are in the same order of magnitude as human localization, reported in [22] to have a mean accuracy of 5 degrees at 30 degrees azimuth (standard deviation 2 degrees).

The planarity scores are also extremely high, with a minimum score of 97.3% (averaged across frequency) for 60 loudspeakers, reducing to 95.5% for 15 loudspeakers. Finally, the control effort is low. It can be seen from the SPL maps that a strong beam is created across the target region. Although such beamforming strongly constrains the localization effect to be within the specified target region, numerous situations are conceivable where this would be desirable (for instance where a single listener is consuming the spatial audio), and the reduction in sound radiation at the room boundaries would be beneficial in terms of room reflections and external sound transmission.

4.2. Comparative performance

The simulations comparing method performance yielded informative results under our evaluation metrics. Examples of the sound fields reproduced at 1 kHz with PP,

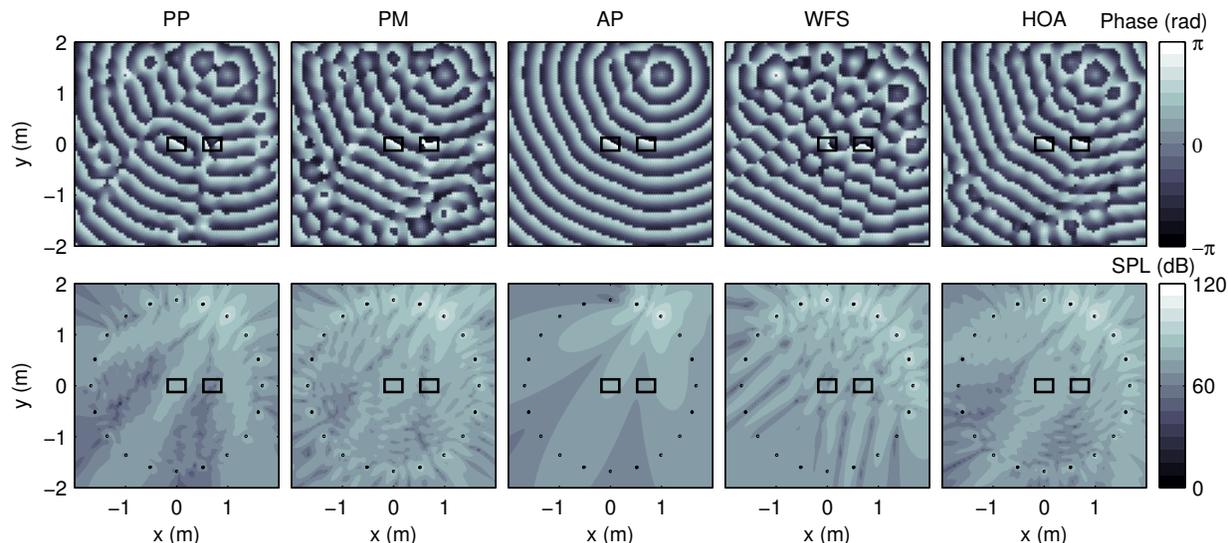


Fig. 4: Phase (top) and sound pressure level (bottom) in the simulated room at 1 kHz, comparing PP, PM, AP, WFS and HOA ($L = 20$), for a listener in the centre zone (A) listening to a virtual source at 30 degrees.

PM, AP, WFS and HOA applied to the 20 channel loudspeaker subset with a far-field virtual source at 30 degrees are shown in Fig. 4. With the reduced number of loudspeakers, the properties of the methods with relatively wide loudspeaker spacing (325 Hz spatial aliasing frequency) and HOA with 10th order expansion, the differences between the methods are clear. The simplest distribution of phase and SPL is given by the AP. The SPL distribution for PP is very similar to this, although by using more loudspeakers, the control effort is reduced. The reproduced sound fields for PM and HOA are also rather similar, with HOA giving a simpler off-centre phase distribution than PM. Finally, WFS is unable to accurately reproduce the target sound field, and complex interference patterns radiate across the listening area. Although these differences between WFS and HOA have been previously reported [12], the similarity between PM and HOA (attempting to exactly reproduce the sound field over a limited area), and PP and AP (attempting to focus the energy to arrive from a certain direction), is an interesting result.

The methods are further compared in Fig. 5, which compares the RMSE, planarity and effort performance over frequency for equally-spaced arrays of 15–60 loudspeakers. Each metric is averaged across 250–6000 Hz and across all 5 virtual source positions. This averaging means that although WFS performed well at lower fre-

quencies, it has a high RMSE and low planarity for each array size due to severe degradations above the spatial aliasing limit. The other methods perform comparably in terms of RMSE and planarity, although the HOA planarity begins to reduce more than PP, PM or AP for 15 loudspeakers. This is the result of the order of expansion beginning to be insufficient for the zone size at high frequencies. PP is always the least-effort approach, with the greatest advantage with the largest loudspeaker array, and alongside PM it has the lowest RMSE (5.5 degrees) for 15 loudspeakers.

Finally, the methods can be compared in terms of the performance in the off-centre listening position (B), both when the source weights were calculated for the central zone (setup A- playback B) and the off-centre zone (B-B). These results, averaged as above, are shown in Fig. 6 for the 20 loudspeaker case. It can be seen that the performance (B-B) in the off-centre zone is slightly worse in terms of RMSE and planarity, and the ranking among methods is similar, to the central listening zone (A-A). It can therefore be concluded that when the listening position is known, the choice of control method is not significantly affected. PP is the least-effort approach in each case, and performs among the best methods for RMSE and planarity. When the source weights calculated for the central zone are evaluated off-centre (A-B), WFS and HOA perform best (although the RMSE is significantly

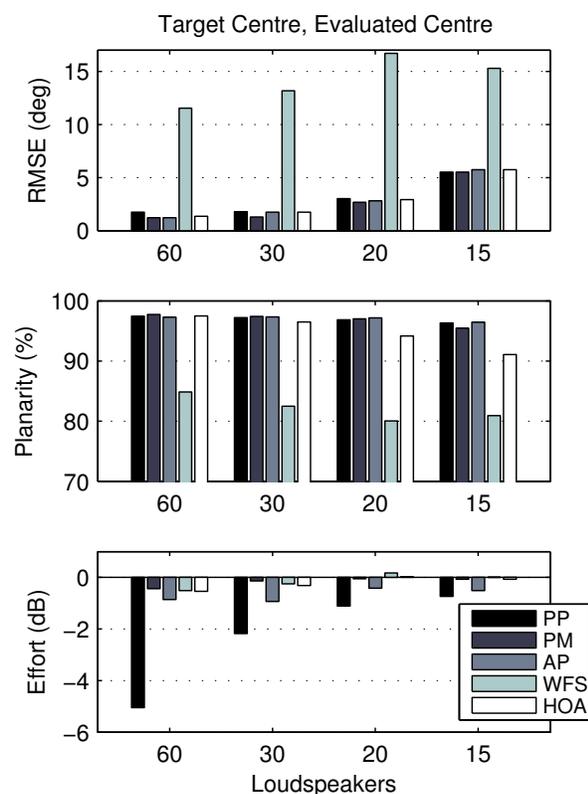


Fig. 5: Performance of PP, PM, AP, WFS and HOA, averaged over frequency and virtual source direction, for decreasing loudspeakers in the circular array, evaluated at the central listening position (A).

increased), which follows from their ability to incorporate the whole reproduction region at lower frequencies. However, the aliasing patterns of WFS (and PM) have a negative effect on the planarity. The present simulation results do not give an indication of how well PP and PM would perform over a larger region in comparison to WFS and HOA. However, it has been verified that PP provides a plausible alternative to these methods for both central and non-central listening positions, when the listening position is known.

5. SUMMARY

The method of planarity panning has been introduced as a means of delivering spatial audio to a listener. The method efficiently focuses the sound energy on the listening region while constraining it to impinge from an angular window about the target direction. Simulations

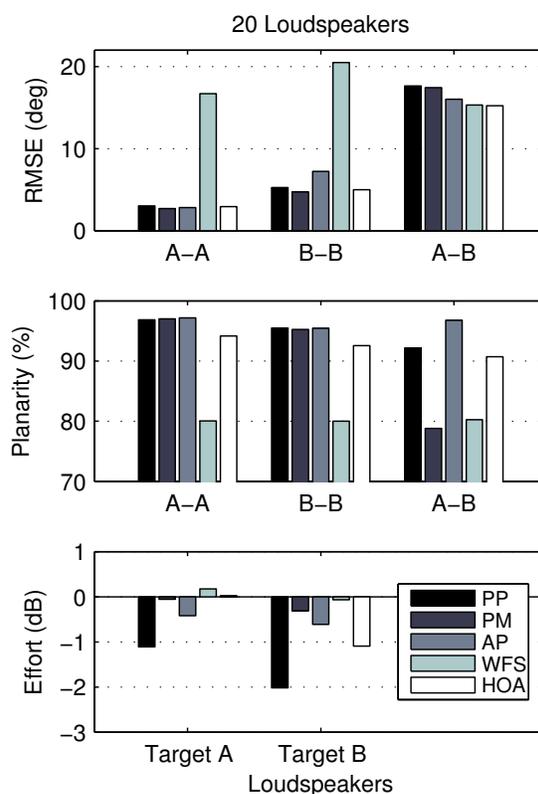


Fig. 6: Performance of PP, PM, AP, WFS and HOA ($L = 20$), averaged over frequency and virtual source direction, evaluated at the central position (A) and non-central position (B), considering each zone as a target (A-A, B-B) and evaluating weights derived for the centre at the off-centre position (A-B).

showed the method to accurately place virtual sources from 5 positions, even when only 15 loudspeakers were used. The method performed comparably with the best of the existing control methods in terms of RMSE and planarity, and was always the least-effort approach. Further work should include perceptual validation of the spatial quality, reproduction in real listening rooms, and application to irregular loudspeaker arrays.

6. ACKNOWLEDGEMENTS

This research has been supported by EPSRC DTG & Grant Ref. EP/L000539/1 (S3A), and extends work undertaken in the POSZ project. Thanks to Marek Olik for contributions to the software used for this work.

7. REFERENCES

- [1] M. A. Gerzon, "Periphony: With-height sound reproduction," *Journal of the Audio Engineering Society*, vol. 21, no. 1, pp. 2–10, 1973.
- [2] J. Ahrens, *Analytic Methods of Sound Field Synthesis*. Berlin: Springer-Verlag, 2012.
- [3] 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.
- [4] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *Proc. 124th AES Convention, Amsterdam, 17-20 May, 2008*.
- [5] O. Kirkeby and P. A. Nelson, "Reproduction of plane wave sound fields," *J. Acoust. Soc. Am.*, vol. 94, no. 5, pp. 2992–3000, 1993.
- [6] M. Poletti, "Robust two-dimensional surround sound reproduction for nonuniform loudspeaker layouts," *J. Audio Eng. Soc.*, vol. 55, no. 7/8, pp. 598–610, 2007.
- [7] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466, 1997.
- [8] 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.
- [9] S. Spors, H. Buchner, R. Rabenstein, and W. Herboldt, "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.
- [10] J. Lopez, A. Gonzalez, and L. Fuster, "Room compensation in wave field synthesis by means of multichannel inversion," in *IEEE Works. Appl. Sig. Proc. Audio Acoust., New Paltz, NY, 16-19 October 2005*, pp. 146–149, 2005.
- [11] T. Betlehem and T. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," *J. Acoust. Soc. Am.*, vol. 117, pp. 2100–2111, 2005.
- [12] S. Spors and J. Ahrens, "A comparison of wave field synthesis and higher-order ambisonics with respect to physical properties and spatial sampling," in *Proc. 125th AES Convention, San Francisco, CA, 2-5 October 2008*, Engineering Society, 2008.
- [13] P. Coleman, P. J. B. Jackson, M. Olik, M. Møller, M. Olsen, and J. A. Pedersen, "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array," *J. Acoust. Soc. Am.*, vol. 135, no. 4, 2014.
- [14] P. Coleman, P. J. B. Jackson, M. Olik, and J. A. Pedersen, "Optimizing the planarity of sound zones," in *Proc. 52nd AES Int. Conf. Sound Field Control, Guildford, UK, 2-4 September, 2013*.
- [15] J.-H. Chang, J.-W. Choi, and Y.-H. Kim, "A plane wave generation method by wave number domain point focusing," *J. Acoust. Soc. Am.*, vol. 128, no. 5, pp. 2758–2767, 2010.
- [16] E. G. Williams, *Fourier Acoustics*. London: Academic Press, 1999.
- [17] Y. Wu and T. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," *IEEE Trans. Audio, Speech Lang. Proc.*, vol. 19, no. 6, pp. 1711–1720, 2011.
- [18] J.-H. Chang, J.-W. Choi, and Y.-H. Kim, "Reproduction of a plane-wave sound field based on wave number domain focusing: Comparison with ambisonics and wfs," in *8th Int. Conf. Motion Vib. Control, Daejeon, Korea, 27-30 August, 2006*.
- [19] J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *J. Acoust. Soc. Am.*, vol. 111, no. 4, pp. 1695–1700, 2002.
- [20] P. J. B. Jackson, F. Jacobsen, P. Coleman, and J. A. Pedersen, "Sound field planarity characterized by superdirective beamforming," in *Proceedings of Meetings on Acoustics*, vol. 19, p. 055056, 2013.
- [21] H. Wierstorf and S. Spors, "Sound Field Synthesis Toolbox," in *132nd AES Convention, Budapest, 26–29 April 2012, 2012*.
- [22] J. C. Makous and J. C. Middlebrooks, "Two-dimensional sound localization by human listeners," *J. Acoust. Soc. Am.*, vol. 87, no. 5, pp. 2188–2200, 1990.