

Sound field planarity characterized by superdirective beamforming

Philip J. B. Jackson^a, Finn Jacobsen^b, Philip D. Coleman^a and Jan Abildgaard Pedersen^c

^a CVSSP, Department of Electronic Engineering, University of Surrey, UK.

^b Acoustic Technology, Dept. of Electronic Engineering, DTU, Denmark.

^c Bang & Olufsen a/s, Peter Bangs Vej 15, DK-7600, Struer, Denmark.

22 Jan. 2013

Abstract

The ability to replicate a plane wave represents an essential element of spatial sound field reproduction. In sound field synthesis, the desired field is often formulated as a plane wave and the error minimized; for other sound field control methods, the energy density or energy ratio is maximized. In all cases and further to the reproduction error, it is informative to characterize how planar the resultant sound field is. This paper presents a method for quantifying a region's acoustic planarity by superdirective beamforming with an array of microphones, which analyzes the azimuthal distribution of impinging waves and hence derives the planarity. Estimates are obtained for a variety of simulated sound field types, tested with respect to array orientation, wavenumber, and number of microphones. A range of microphone configurations is examined. Results are compared with delay-and-sum beamforming, which is equivalent to spatial Fourier decomposition. The superdirective beamformer provides better characterization of sound fields, and is effective with a moderate number of omni-directional microphones over a broad frequency range. Practical investigation of planarity estimation in real sound fields is needed to demonstrate its validity as a physical sound field evaluation measure.

INTRODUCTION

The reproduction of sound fields that may be perceived uniformly as coming from a specified direction is an essential component of modern spatial audio applications where the quality of listening experience is driving progress [1, 2, 3]. One application of multi-channel audio reproduction forms sound zones, within which different content may be presented [4, 5, 6, 7, 8, 9, 10, 11]. Such systems can produce interfering sound fields with fluctuating sound level, particularly those based on energy minimization, e.g., [6, 12, 13], which can be detected using an appropriate microphone array. It is desirable, therefore, to derive an interpretation of the microphone signals that provides an objective measure of the reproduction quality across the zone, ideally at each frequency of interest.

The reproduction error (root-mean-square (rms) of the differences between target and reproduced pressure) provides a useful summary statistic of the scale of the deviations, including all possible errors and artifacts, e.g., due to gain and phase variation and misalignment, and waves arriving from various directions [4, 5, 7, 11]. Another measure is acoustic contrast [6, 11], the ratio of energy between two regions based on spatially-averaged squared sound pressures. Again, ignoring sound velocity, it fails to distinguish between spatial modifications within zones and does not help in diagnosis of the output. Neither of these measures captures the phase relationships between target and reproduced sound fields or considers the directions of sound propagation, so further descriptors are needed to differentiate.

Expressing the reproduced sound field in plane wave terms specifies the objective, and gives a valuable alternative perspective [8]. Thus, the similarity to a plane wave sound field, or *planarity*, is proposed. Natural, localized sources in the free, far field approximate plane waves, whereas the consistency of interaural differences used in spatial perception can be destroyed by the interference patterns from multiple sources [14, 2]. Planarity allows the consistency with a plane wave to be evaluated, without presupposing an orientation or propagation delay.

Spatial audio coding too has an interest in the direction of sound propagation. Directional audio coding (DirAC) uses, for example, first-order (B-format) microphones to determine source energy and intensity, and derive the diffuseness [15]. Rather than coincident pressure and velocity microphones, Wiederhold *et al.* estimated sound intensity from four pressure microphones mounted on a sphere [16]. Using a greater number of sample points offers improved robustness and spatial resolution over point estimates of sound intensity, which can be sensitive to calibration errors [17, 18, 19]. Choi and Kim used a microphone grid to extract 1st-order intensity estimates over a target region [20]. Khaykin and Rafaely tested a high-order spherical array for spatial filtering of room impulse responses using conventional and superdirective beamforming methods [21], showing better performance with the latter.

Here, a measure of planarity is proposed that relates the amount of sound energy traveling in one direction (that of the principal plane wave component) to the total energy flux. This study is restricted to waves in 2D but offers straightforward extension to 3D. The procedure includes estimation of the distribution of energy flux with respect to its direction of arrival (DOA) from pressure transducers (omni-directional microphones). The square grid of microphones acts as the primary basis for our analysis (see Fig. 1), reflecting common practice in sound field sampling. Two planarity estimation methods are examined and compared through simulations: the delay-and-sum beamformer (which is equivalent to spatial Fourier decomposition), and the acoustic-contrast beamformer representing the state-of-the-art in optimal array processing. The efficacy with various microphone array configurations is tested, revealing notable differences in resolution, performance, consistency and effective frequency range with a limited microphone spacing.

In the next section, the method of planarity estimation is presented for the conventional and superdirective beamformers. Simulations in Section 3 provide comparisons over frequency, number of microphones, array orientation and various array configurations. Section 4 concludes.

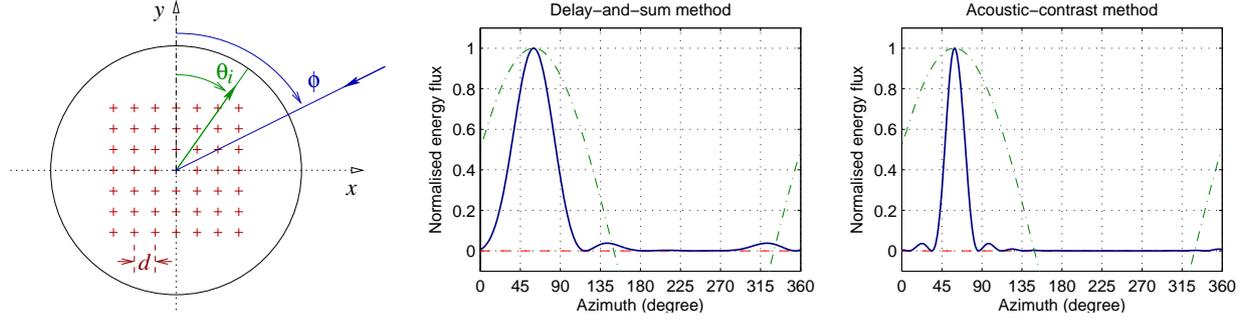


FIGURE 1: (a) Diagram of the cartesian reference frame (open arrows with dotted lines), direction of arrival of the principal plane wave ϕ (inward arrow), default microphone array (crosses) with $n_x = n_y = 7$ and $n_M = 49$, steering angle θ_i (outward arrow), and inter-element spacing d . Normalized energy flux distributions (solid thick lines) for (b) delay-and-sum and (c) acoustic-contrast beamforming; thin dash-dot lines show cosine weighting of the intensity estimate aligned to the principal plane wave, $\phi = 60^\circ$, $f = 1$ kHz, $n_M = 49$.

THEORY AND METHODS

This section sets planarity in the context of sound intensity and energy flux. Two methods of using omni-directional microphones are presented for characterizing the sound energy distribution over DOA and obtaining planarity estimates: delay-and-sum and superdirective beamforming.

A sound field may be represented as a sum of plane waves, each with an amplitude and DOA that determine the amount and direction of energy propagation [8]. The magnitude of intensity in any specified direction ϕ (with unit vector $\mathbf{u}_\phi = [\sin\phi \ \cos\phi]^T$) is the sum of all wave components in that direction, $I_\phi = \mathbf{I} \cdot \mathbf{u}_\phi = \sum_i w_i \mathbf{u}_i \cdot \mathbf{u}_\phi$ where \mathbf{I} is the sound intensity vector, \mathbf{u}_i the unit vector associated with the i th component's DOA, and \cdot denotes the inner (dot) product. The energy flux distribution over DOA ($0 \leq \theta_i < 2\pi$) is $w_i = \psi_i^* \psi_i / 2$, $\forall i = 1..n_\Theta$, where ψ_i is the i th wave, n_Θ the number of angles, and $*$ denotes complex conjugate. For plane waves, the product of the energy density and rate of propagation gives the total energy flux, $E c = \sum_i w_i$. Planarity, η , is defined as the ratio between the intensity component in the direction of the sound field's principal plane wave $\hat{i} = \arg \max_i w_i$ and the total energy flux:

$$\eta = \frac{\mathbf{I} \cdot \mathbf{u}_{\hat{i}}}{E c} = \frac{\sum_i w_i \mathbf{u}_i \cdot \mathbf{u}_{\hat{i}}}{\sum_i w_i}. \quad (1)$$

The planarity scores lie in the range $-1 < \eta \leq 1$, although negative scores only occur when multiple secondary components counter the principal component. A plane wave gives $\eta = 1$; a standing wave or perfectly diffuse field (i.e., where $\lim_{t \rightarrow \infty} \langle \mathbf{I} \rangle_t = 0$) gives $\eta = 0$. Planarity quantifies the proportion of sound energy coming from the principal direction, and therefore acts as a measure of a sound field's spatial alignment. Both methods below use Eq. (1) to compute η , but differ in their energy flux distribution estimates, w_i . Hereafter, the general expressions are applied to harmonic signals, and complex Fourier coefficients employed.

Delay-and-sum beamforming steers the microphone array to 'look' along each candidate azimuth, find the DOA of maximum energy, then calculate the planarity for this direction. As input, the method takes the wavenumber k , number of microphones n_M , their positions \mathbf{r}_s and complex pressures $p_s \ \forall s \in 1..n_M$. The look directions were distributed regularly with $n_\Theta = 360$.

First, the n_Θ -by- n_M steering matrix $\mathbf{H} = \{h_{i,s}\}$ is specified, with elements for each look direction i and microphone s , $h_{i,s} = e^{-jk\mathbf{r}_s \cdot \mathbf{u}_i} / n_M \ \forall i = 1..n_\Theta, s = 1..n_M$. By applying the steering matrix to the set of microphone pressures $\mathbf{p} = [p_1 \dots p_{n_M}]^T$, the vector of steered responses is obtained:

$$\boldsymbol{\psi} = [\psi_1 \dots \psi_{n_\Theta}]^T = \mathbf{H} \mathbf{p}. \quad (2)$$

Hence, the energy flux distribution $w_i = \psi_i^* \psi_i / 2$ is computed, the global maximum $\hat{i} = \arg \max_i w_i$ found, and planarity estimated as Eq. (1).

The delay-and-sum beamformer directly estimates ψ_i , which equates to the indirect estimate by DFT and interpolation via spatial Fourier decomposition evaluated at points on the wavenumber circle. Figure 1b shows the energy flux distribution obtained with the delay-and-sum beamformer for a plane wave with DOA $\phi = 60^\circ$, frequency $f = 1$ kHz, and $n_M = 49$ microphones. There is a broad energy peak at $\theta_i = 60^\circ$ and side lobes at 140° and 320° .

Superdirective beamforming adjusts the steering matrix to optimize the array's direction resolution giving a refined estimate of the energy flux distribution, which is crucial for calculating planarity. The acoustic contrast method was chosen, which seeks to maximize the ratio of energy from the pass band of DOAs to that of the stop range [6, 22]. Rather than applying the method to loudspeakers [23, 24], here it is reciprocally applied to microphones.

Steering weights are found by eigenvalue decomposition of the squared pass-beam matrix multiplied by the inverse squared stop-range matrix. First, the microphone responses for each look direction are defined based on the plane wave Green's function, $g_{i,s} = e^{jk\mathbf{r}_s \cdot \mathbf{u}_i / n_M} \forall i = 1..n_\Theta, s = 1..n_M$ then grouped into pass beam \mathbf{B} and stop range \mathbf{D} :

$$\begin{aligned} \mathbf{B}_i &= \{g_{i',s}\} & \forall s \in 1..n_M, i' \in i \pm \delta_P \\ \mathbf{D}_i &= \{g_{i'',s}\} & \forall s \in 1..n_M, i'' \notin i \pm \delta_S \end{aligned} \quad (3)$$

The steering vector for each look direction \mathbf{h}_i is the eigenvector that corresponds to the maximum eigenvalue, $\max \lambda$, with β regularizing the inversion:

$$\mathbf{h}_i = \arg \max \text{eig} \left\{ \left(\mathbf{D}_i^H \mathbf{D}_i + \beta I \right)^{-1} \mathbf{B}_i^H \mathbf{B}_i \right\} \quad (4)$$

These are collated to form the steering matrix, $\mathbf{H} = \left[\mathbf{h}_1 \dots \mathbf{h}_{n_\Theta} \right]^T$, which is then applied as in Eq. (2). From here onwards, the two methods are identical.

The choice of beam width and stop range depends on the desired trade off between sharpness of the main beam and suppression of side lobes, considering array resolution and robustness. For each look direction, a pass beam $\delta_P = \pm 3^\circ$ on either side of the center angle was included; the stop range covered all other angles outside $\delta_S = \pm 6^\circ$. These settings gave a reasonable compromise. The matrix inversion had Tikhonov regularization to mitigate numerical problems associated with singularity and to restrict the array's white noise gain [22], A suitable regularization of $\beta = 10^{-4}$ was determined empirically, though the response was not sensitive to this choice.

By taking the response at other azimuths into account, the superdirective beamformer suppresses the side lobes and sharpens the energy distribution optimally. In Fig. 1c, the energy flux distribution shows less than half the beamwidth compared with delay-and-sum beamforming. The side lobes are much closer in azimuth to the true DOA and contain significantly less energy, improving the accuracy of the planarity estimates on both counts. The following section provides an evaluation of these methods.

SIMULATIONS UNDER ANECHOIC CONDITIONS

Evaluation of the behavior of the proposed estimation methods was conducted by simulations of ideal sound fields and microphones in the free field. The air was assumed to support isentropic propagation with sound speed, $c = 340$ m/s, the microphones to be noiseless omni-directional point sensors with unit gain, and each sound field to comprise ideal plane wave or point-monopole sources. A diffuse field was simulated by the superposition of 60 regularly-spaced, equal-amplitude plane waves with random phase, similar to [25, 26, 27]. These conditions enable direct examination of the methods' fundamental properties, and indicate bounds on practical performance.

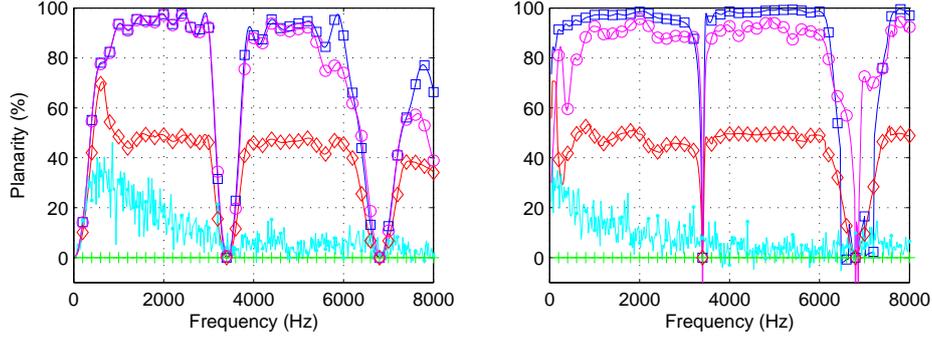


FIGURE 2: Planarity estimates with respect to source frequency f ($\phi=0^\circ$, $n_M=49$), by (a) delay-and-sum and (b) acoustic-contrast beamforming: (square) plane, (+) standing, (\diamond) check, (\cdot) diffuse and (\circ) point source. Diffuse-field results are means, $N=6$.

To test a range of planarity scores with a value in the middle, an indication of the measure's statistical variability, and an illustration of the effect of source displacement from the array, five sound field types were tested [28, 29]: 1. *planar*, a single plane wave; 2. *standing*, two equal-and-opposite plane waves; 3. *check*, two plane waves 90° apart; 4. *diffuse*, random plane waves from all directions; 5. *point*, a point source 1 m from array center. The nominal planarity for the planar field is 100%; for the standing field, it is 0% since the energy flux cancels. The check field has a nominal value of 50% for whichever of the two components at 0° and 90° is labeled as principal, since the other is orthogonal. In the diffuse field, the pressure and velocity at any point are independent random variables with finite mean magnitude and variance [25, 28]. All DOAs are equally likely, so the mean intensity is zero, although the intensity's mean magnitude, $\mathcal{E}\{|\mathbf{I}|\}$, is finite [28]. A distinction emerges between $\mathcal{E}\{\mathbf{I}\cdot\mathbf{u}_i\}$ and $\mathcal{E}\{|\mathbf{I}|\}$. For a fully-resolved diffuse field, \mathbf{u}_i is uncorrelated with \mathbf{I} , so $\mathcal{E}\{|\eta|\}=0$ and the nominal planarity is 0%. Two factors reduce the planarity of the point field in relation to the planar field: SPL deviation and wavefront curvature. These increase as the source approaches the array or the array grows, but do not depend on frequency, so a nominal value of $\sim 90\%$ is expected with our default settings. All simulations were performed with microphone spacing of $d=0.05$ m (spatial aliasing frequency, 3.4 kHz) and implemented in Matlab.

Source azimuth. To study the effect of incidence angle on the array, the true DOA of the principal plane wave component was varied in 1-degree azimuth intervals, $\phi=1^\circ..360^\circ$. Planarity scores were computed for the five field types with both beamforming methods at $f=1$ kHz (not plotted). At this frequency, the delay-and-sum results sat around 91% for the plane wave source. Better estimates were obtained in directions for which either rows or columns of the array were in broadside/endfire configuration. The standing wave consistently produced values equal to zero. The check sound field created from two orthogonal plane wave components gave planarity estimates just below 50%. The diffuse-field estimates varied widely about a mean of 27%, including values above 70% and as low as -10% . The point source obtained values less than 1% from those of the plane wave, which highlights the insensitivity of conventional beamforming (and spatial Fourier decomposition) to small changes in the wavefronts in the array region.

In comparison, superdirective beamforming showed improved discernment of the test cases. The plane field scored 97%; the check mean rose from 47% to 48%; the diffuse readings ranged less widely from -10% to below 50% with lower mean of 15%; the point source's planarity score also increased to an average 92% but generally not as much as that of the plane wave. The separation between the plane and point sources was significant for most azimuths but disappeared at diagonal DOAs, $\phi=\{45^\circ, 135^\circ, 225^\circ, 315^\circ\}$.

Frequency. To examine the relationship of inter-element spacing to the wavelength, the performance over frequency was studied ($\phi=0^\circ$). In the present conditions, varying the frequency is equivalent to varying the inter-element spacing, since the response depends on the ratio of any

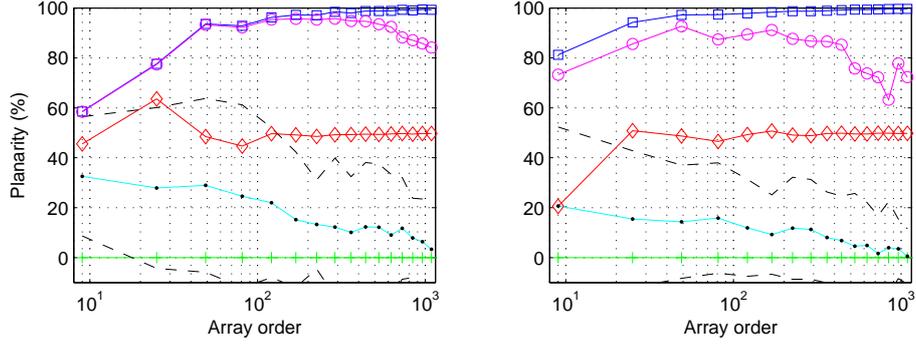


FIGURE 3: Planarity estimates with respect to number of microphones n_M ($\phi = 0^\circ$, $f = 1$ kHz), by (a) delay-and-sum and (b) acoustic-contrast beamforming: (square) plane, (+) standing, (◊) check, (·) diffuse and (○) point source. Diffuse-field means shown with ± 2 standard deviations (dashed), $N = 30$.

distances to the wavelength. Results are presented for $50 \text{ Hz} \leq f \leq 8 \text{ kHz}$.

Figure 2a shows the marked impact of frequency variation on the performance of the array, depends heavily on the relationship between the microphone configuration and the wavelength. The key dimensions are the inter-element spacing $d = 0.05$ m and the array aperture $d n_x = d n_y = 0.35$ m, whose spatial aliasing frequencies (3.4 kHz and 0.5 kHz, respectively) effectively bound the operating range of the conventional beamformer. Inside these bounds, a similar pattern of results was obtained as those testing azimuth, plus there was a tendency for the diffuse-field scores to decline over frequency and become less variable. The pattern of behaviour was repeated for aliased frequencies, $3.4 \text{ kHz} \leq f \leq 6.8 \text{ kHz}$ and again above that, while above 5 kHz the scores decreased for plane, check and especially point source fields. The reason for the rise in diffuse scores at low frequencies is linked to the wavelength becoming large in comparison with the spacing of the microphones, which reduced the independence of their sound pressures in the diffuse field [30]. As the array’s spatial resolution degrades, \mathbf{u}_i becomes positively correlated with \mathbf{I} leading to a significant positive bias on η scores, until ultimately the energy flux is completely unresolved and the planarity estimates for all field types tend to zero.

The results with the superdirective beamformer (Fig. 2b) have a similar overall structure, yet once again its scores tended to be higher and exhibit greater discriminability. This can be seen in the increase of scores in the planar field, the separation between scores in planar and point fields, and the lower diffuse field scores, especially in the extended low frequency range, 0.2 kHz–1.0 kHz. Note too the narrower range of planarity scores in the diffuse field, which tend to decrease as frequency increases and the wavelength decreases. The roll off in performance at low frequency and in the vicinity of aliasing frequencies affected a narrower band than with the delay-and-sum beamformer. At these aliasing-frequency dips, however, the rank order can be disrupted, e.g., the point source planarity can score above that of the plane wave sound field (see 6.5 kHz–7.3 kHz).

Array order. The accuracy of the beamformer is directly linked to the number of microphones, as higher-order arrays more effectively suppress side lobes and the effects of spatial aliasing, and larger arrays achieve finer spatial resolution. For comparison, simulations were conducted for square regularly-spaced arrays of varying size, $n_M = \{3^2, 5^2, 7^2, \dots, 33^2\}$. The results are shown in Fig. 3.

For the smallest array ($n_M = 9$), the array aperture was too small adequately to resolve the 0.34-m wavelength plane waves ($f = 1$ kHz). For the delay-and-sum beamformer (Fig. 3a), the plane field’s planarity first exceeded 90% with $n_M = 49$; whereas this occurred with $n_M = 25$ for the superdirective beamformer. Once again, the same trends were observed, including a fall in planarity of the diffuse field (mean and variance) as the number of microphones was increased, and better separation of the plane and point source fields. The scores approached asymptotic

<i>Array</i>	<i>Method</i>	planar (100%)	standing (0%)	check (50%)	diffuse (0%)	point (~90%)
Cross	DS	70.0 ±5.1	0.0 ±0.0	36.8 ±9.2	18.1 ±27.8	67.7 ±14.6
	AC	60.1 ±13.8	0.0 ±0.0	32.1 ±14.8	7.9 ±20.8	27.8 ±20.1
Circle	DS	65.4 ±0.0	0.0 ±0.0	12.9 ±0.0	16.2 ±24.1	60.3 ±0.0
	AC	98.8 ±0.0	0.0 ±0.0	48.8 ±0.0	8.1 ±20.1	77.3 ±0.0
Dual	DS	76.3 ±0.0	0.0 ±0.0	30.0 ±0.0	20.9 ±27.1	75.1 ±0.0
	AC	97.6 ±0.0	0.0 ±0.0	46.8 ±0.0	13.5 ±24.6	94.0 ±0.0
Square	DS	90.6 ±4.9	0.0 ±0.0	46.6 ±3.9	27.7 ±30.4	90.1 ±4.9
	AC	97.0 ±0.2	0.0 ±0.0	48.1 ±1.1	15.9 ±24.7	92.4 ±4.7
Hex	DS	90.5 ±0.9	0.0 ±0.0	49.4 ±0.8	27.8 ±33.6	90.1 ±0.9
	AC	95.1 ±0.7	0.0 ±0.0	54.6 ±4.2	17.0 ±25.2	81.9 ±6.6

TABLE 1: Planarity estimates for each field type using various array geometries ($f = 1$ kHz, $n_M = 49$), by delay-and-sum (DS) and acoustic-contrast (AC) beamforming; mean across DOA with ± 2 standard deviations, $N = 180$.

values for large n_M , e.g., 100 % for plane, 50 % for check, and 0 % for standing and diffuse.

Array geometry. The final set of simulations investigated the performance of a variety of array geometries for planarity estimation. One alternative to a regular grid comes from the perspective of the Kirchhoff-Helmholtz integral over an enclosing surface or, in the 2D case, around a surrounding perimeter. For a circular sound zone, one would configure microphones as evenly-spaced discrete samples on a circle. A set of n_M microphones would theoretically capture n_M spatial modes, subject to spatial aliasing. Yet, n_M modes can be estimated by any arbitrary set of n_M independent measurements. In a diffuse field, the spatial coherence between microphones reaches a minimum at a spacing equal to half the wavelength, i.e., at the spatial aliasing frequency [30], which therefore determines one essential design parameter. Smaller separations than this erode the independence of the measurements [31]. So, maintaining independence over a wide range of frequencies requires that the sample points be well separated. For the fixed sample area of a sound zone and a given number of microphones, the optimal spacing is therefore given by a regular grid. By packing microphones in a hexagonal grid, a $\sim 15\%$ increase in the number of microphones can be attained relative to a rectangular grid; conversely, for fixed array order and microphone spacing, the hexagonal arrangement yields the most compact configuration.

This subsection examines the effect of array geometry on the performance of delay-and-sum and superdirective planarity estimates, using five configurations with the same number of microphones and minimum spacing, $n_M = 49$ and $d = 0.05$ m (diameter of enclosing circle):

1. *Cross* (1.20 m), four 12-element lines emanating from a central element
2. *Circle* (0.78 m), all 49 elements around the circumference of a single circle
3. *Dual* (0.44 m), elements on two concentric circles: 21 on the inner circle, 28 on the outer one
4. *Square* (0.42 m), the 7×7 rectangular grid used previously
5. *Hex* (0.36 m), the 49 points on a hexagonal grid closest to the center point

The results for azimuths at $f = 1$ kHz are summarized in Table 1. In all cases except Cross, the order of the plane, point, check and standing waves was the same and as expected. In the diffuse field, considerable variability occurred, so that the planarity could exceed that in the check field or fall below those of the standing waves at zero. The diffuse-field mean sat between that of the standing and check waves in all cases but Circle with delay-and-sum estimation, for which the check scores were a few percentage points below the diffuse-field mean.

In the delay-and-sum case, Square and Hex showed comparable performance with generally higher planarity scores for the plane field than Dual and Circle. However, they did not adequately distinguish between the plane wave and point source. Cross, Dual and Circle also gave correspondingly lower mean and variance for the diffuse field with their compressed scales. In addition, Cross performance was variable over azimuth.

In the superdirective case, the planarity of the plane wave field was markedly higher across all geometries, except Cross. The accuracy of the planarity score increased monotonically across these geometries with the span of the array, both in terms of absolute value and variability with DOA. Conversely, the score for the diffuse field was increasingly suppressed as the array grew. All geometries showed discrimination between the plane and point source, although the Square failed to do so for diagonal directions. Hex gave comparable values for the point source, on average, while only about half the extent of Circle. Cross performance was generally poor.

In summary, results showed that superdirective beamforming enhanced the performance of all array geometries for purposes of planarity estimation. For the mid-frequency range 0.5 kHz–3.0 kHz, the Dual array with acoustic-contrast beamforming gave the most consistently accurate results. Hex was the most compact array and still provided adequate separation of the different sound field types. On the other hand, Circle yielded the highest absolute planarity scores across a broad range, 0.2 kHz–4.0 kHz. This is consistent with findings of Hulsebos *et al.* who compared the performance of line, cross and circle configurations for wave-field synthesis sound reproduction [32]. Square offered a practical compromise between size and performance over the range 0.5 kHz–6.2 kHz, except around the 3.4-kHz spatial aliasing frequency. Evidently, the choice of geometry is crucial in determining an array's performance. By further investigation of alternative geometries, the configuration could be tailored to give the best combination of accuracy, consistency and compactness for a specific application.

CONCLUSIONS

Two methods were presented for estimating the planarity of a sound field over a region where omni-directional microphones are placed, delay-and-sum and superdirective beamforming. A series of experimental simulations was conducted to investigate their performance under ideal conditions. The effects of orientation of the array with respect to the source's DOA, the frequency, the array order and geometry were studied to establish limitations in the effective operation and specification of the microphone-array-based estimation of sound intensity. The superdirective (acoustic-contrast) beamforming method generally showed superior performance in terms of the accuracy of the planarity estimates, their variability and operating range. Planarity can augment the set of evaluation measures applied in sound field control and spatial audio reproduction. Further work could investigate the potential relationship between listeners' perception of target sound quality and planarity in real sound fields, and assess the methods' robustness to errors.

ACKNOWLEDGMENTS

This research is linked with the POSZ project (www.posz.org) supported by B&O and EPSRC.

REFERENCES

- [1] J. Engdegaard, B. Resch, C. Falch *et al.*, "Spatial audio object coding (SAOC) – the upcoming MPEG standard on parametric object based audio coding", in *Proc. 124th AES Conv., Amsterdam*, 7377 (2008).
- [2] J. Breebaart and C. Faller, *Spatial Audio Processing: MPEG Surround and Other Applications* (John Wiley & Sons, Chichester, UK) (2008).
- [3] F. Rumsey, *Spatial Audio*, Music Technology Series, 2nd ed. (Taylor & Francis, Oxford, UK) (2012).
- [4] O. Kirkeby and P. A. Nelson, "Reproduction of plane wave sound fields", *JASA* **94**, 2992–3000 (1993).
- [5] D. Ward and T. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers", *IEEE Trans. Speech & Audio Process.* **9**, 697–707 (2001).
- [6] J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources", *JASA* **111**, 1695–1700 (2002).

- [7] T. Betlehem and T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms", *JASA* **117**, 2100–2111 (2005).
- [8] S. Spors, H. Buchner, R. Rabenstein, and W. Herbordt, "Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering", *JASA* **122**, 354–369 (2007).
- [9] Y. Wu and T. Abhayapala, "Theory and design of soundfield reproduction using continuous loudspeaker concept", *IEEE Trans. Audio, Speech, Lang. Process.* **17**, 107–116 (2009).
- [10] M. Poletti, F. Fazi, and P. Nelson, "Sound-field reproduction systems using fixed-directivity loudspeakers", *JASA* **127**, 3590–3601 (2010).
- [11] J.-H. Chang and F. Jacobsen, "Sound field control with a circular double-layer array of loudspeakers", *JASA* **131**, 4518–4525 (2012).
- [12] S. Elliott, J. Cheer, H. Murfet, and K. Holland, "Minimally radiating sources for personal audio", *JASA* **128**, 1721–1728 (2010).
- [13] M. Shin, S. Lee, F. Fazi, P. Nelson, D. Kim, S. Wang, K. Park, and J. Seo, "Maximization of acoustic energy difference between two spaces", *JASA* **128**, 121–131 (2010).
- [14] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, 2nd ed. (MIT Press, Cambridge MA) (1996).
- [15] J. Vilkamo, T. Lokki, and V. Pulkki, "Directional audio coding: Virtual microphone-based synthesis and subjective evaluation", *J. Audio Eng. Soc.* **57**, 709–724 (2009).
- [16] C. Wiederhold, K. Gee, J. Blotter, and S. Sommerfeldt, "Comparison of methods for processing acoustic intensity from orthogonal multimicrophone probes", *JASA* **131**, 2841–2852 (2012).
- [17] F. Jacobsen and H.-E. de Bree, "A comparison of two different sound intensity measurement principles", *JASA* **118**, 1510–1517 (2005).
- [18] B. Gover, J. Ryan, and M. Stinson, "Microphone array measurement system for analysis of directional and spatial variations of sound fields", *JASA* **112**, 1980–1991 (2002).
- [19] H. Silverman, W. Patterson III, and J. Sachar, "Factors affecting the performance of large-aperture microphone arrays", *JASA* **111**, 2140–2157 (2002).
- [20] J.-W. Choi and Y.-H. Kim, "Manipulation of sound intensity within a selected region using multiple sources", *JASA* **116**, 843–852 (2004).
- [21] D. Khaykin and B. Rafaely, "Acoustic analysis by spherical microphone array processing of room impulse responses", *JASA* **132**, 261–270 (2012).
- [22] S. Elliott, J. Cheer, J.-W. Choi, and Y. Kim, "Robustness and regularization of personal audio systems", *IEEE Trans. Audio, Speech, Lang. Process.* **20**, 2123–2133 (2012).
- [23] J.-W. Choi, Y. Kim, S. Ko, and J. Kim, "Super-directive loudspeaker array for the generation of personal sound zone", in *Proc. 125th AES Conv., San Francisco, CA*, 7620 (2008).
- [24] J.-W. Choi, Y. Kim, S. Ko, and J. Kim, "A differential approach for the implementation of superdirective loudspeaker array", in *Proc. 128th AES Conv., London*, 8032 (2010).
- [25] F. Jacobsen, "The diffuse sound field", Technical Report 27, Acoustics Laboratory, Technical University of Denmark, Lyngby (1979).
- [26] B. Rafaely, "Spatial-temporal correlation of a diffuse sound field", *JASA* **107**, 3254–3258 (2000).
- [27] M. Kuster, "Spatial correlation and coherence in reverberant acoustic fields: Extension to microphones with arbitrary first-order directivity", *JASA* **123**, 154–162 (2008).
- [28] F. Jacobsen, "Active and reactive sound intensity in a reverberant sound field", *J. Sound & Vib.* **143**, 231–240 (1990).
- [29] F. Fahy, *Sound Intensity*, 2nd ed. (Taylor & Francis, Oxford, UK) (1995).
- [30] F. Jacobsen and T. Roisin, "The coherence of reverberant sound fields", *JASA* **108**, 204–210 (2000).
- [31] D. Lubman, "Spatial averaging in a diffuse sound field", *JASA* **46**, 532–534 (1969).
- [32] E. Hulsebos, D. de Vries, and E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by wave-field synthesis", *J. Audio Eng. Soc.* **50**, 779–790 (2002).