A Measure Based on Beamforming Power for Evaluation of Sound Field Reproduction Performance

Ji-Ho Chang [1],* and Cheol-Ho Jeong [2]

[1]   Center for Fluid Flow & Acoustics, Division of Physical Metrology,
    Korea Research Institute of Standards and Science, Daejeon 34113, Korea
[2]   Acoustic Technology, Department of Electrical Engineering, Technical University of Denmark,
    DK-2800 Kongens Lyngby, Denmark; chj@elektro.dtu.dk
*   Correspondence: chang.jiho@gmail.com; Tel.: +82-42-868-5309

Abstract: This paper proposes a measure to evaluate sound field reproduction systems with an array of loudspeakers. The spatially-averaged squared error of the sound pressure between the desired and the reproduced field, namely the spatial error, has been widely used, which has considerable problems in two conditions. First, in non-anechoic conditions, room reflections substantially deteriorate the spatial error, although these room reflections affect human localization to a lesser degree. Second, for 2.5-dimensional reproduction of spherical waves, the spatial error increases consistently due to the difference in the amplitude decay rate, whereas the degradation of human localization performance is limited. The measure proposed in this study is based on the beamforming powers of the desired and the reproduced fields. Simulation and experimental results show that the proposed measure is less sensitive to room reflections and the amplitude decay than the spatial error, which is likely to agree better with the human perception of source localization.

Keywords: sound field reproduction; higher-order ambisonics (HOA); evaluation of reproduction performance; beamforming power; loudspeaker arrays

1. Introduction

Sound field reproduction methods, such as higher-order ambisonics (HOA) [1–8] and wave field synthesis (WFS) [9–11], attempt to reproduce a sound field that is similar both physically and perceptually to what is intended. Multi-zone sound field reproduction methods have the same purpose for each sound zone [12–15]. Sound field reproduction systems using those reproduction methods have been implemented in several places [16–27], and there have been a number of studies on physical validations [17,18,25,26,28–30] and perceptual evaluations [17–20,22,28,30–33] with the reproduction systems. There are spatial [17,18,25,26,28,30], temporal, and spectral features [29] in the physical validations, which are related to perceptual attributes, such as localization, source width, and sound quality. This paper is concerned with a physical evaluation for the spatial feature, which is related to human localization.

For the physical evaluation of the spatial feature, the spatially-averaged squared pressure error between the desired and the reproduced fields, called the spatial error in what follows, has been widely used [5–8,12,14,15,25,34], whilst listening tests are conducted for perceptual evaluation of human localization. However, the spatial error does not always correspond to human localization [29]. For example, if a sound source is reproduced in non-anechoic rooms, room reflections significantly increase the spatial error, although human localization is affected to a lesser degree. In addition, differences in the amplitude decay with the distance increase the spatial error consistently, but play a less important role from a perceptual view point [35–38]. The aim of this paper is to propose a

physical measure to evaluate the spatial feature of the reproduced sound fields that potentially has better agreement with human localization than the spatial error does. Since it is nearly impossible that a physical measure perfectly reflects human localization as long as human perception of source localization is not fully understood, this paper attempts to reduce the effects of the room reflections and amplitude decay with distance while evaluating the directions of sound waves as effectively as the spatial error.

Room reflections and spatial amplitude decay are relevant issues to practical reproduction systems. Most reproduction systems are installed in highly-damped rooms, not in anechoic conditions, although most reproduction methods strictly assume a free-field condition. There are a few studies that attempt to make use of room reflections [39,40], but these methods are feasible only if the room geometry and absorption properties are known, which is not always possible in practice. Moreover, linear arrays and circular arrays are implemented in many cases, which is called a 2.5-dimensional (2.5D) reproduction. When a 2.5D reproduction is used to reproduce a monopole source, the reproduced monopole does not have the amplitude decay of $1/r$, where r is the distance to the monopole. Unwanted reflections and unsuccessful reproduction of the amplitude decay would consistently magnify the spatial error, of which the perceptual effect is found to be rather limited [35–38].

In order to overcome the limitations of the spatial error, the present study proposes a measure based on the beamforming power, or the beam-power. The beam-powers are calculated based on the sound field in a control region, whereas the other binaural measures, such as inter-aural time difference (ITD) or inter-aural level difference (ILD), are specific to the position and direction of the listener at each time. The main reason why beamforming is used for the measure is that it can concentrate the direct sound energy in a control region along its propagating direction, whereas the reflections are, more or less, uniformly distributed. Thus, the effect of the reflections can be suppressed by a directional filter that gives a higher weighting around the intended direction of a reproduced source. After applying the directional filter, the error of the beam-power between the desired and the reproduced fields is quantified and termed the beam-power error. The effect of the difference in the amplitude decay vanishes when the beam-power is normalized (See Section 3). A typical circular loudspeaker array with 2.5D HOA reproduction is used as an example.

The research problems of this paper are defined in Section 2. The proposed measure and its basic properties are investigated via simulations in Section 3. In Section 4, the beam-power error is compared with the spatial error in a simple pre-experiment and experiments with 2.5D HOA with a circular array of 16 loudspeakers in a damped room.

2. Problem Statements

2.1. Spatial Error in Sound Field Reproduction

Figure 1 illustrates a typical sound field reproduction system that consists of multiple loudspeakers in an array. The number of the loudspeakers is denoted as L, and the position of the lth loudspeaker is denoted as $\vec{r}_s^{(l)}$. The reproduced sound pressure at an arbitrary point in a room $\vec{r}$ can be described as:

$$P\left(\vec{r};\omega\right) = \sum_{l=1}^{L} H^{(l)}\left(\vec{r}\,\middle|\,\vec{r}_s^{(l)};\omega\right) Q^{(l)}(\omega)\, S(\omega), \tag{1}$$

where $H^{(l)}\left(\vec{r}\,\middle|\,\vec{r}_s^{(l)};\omega\right)$ is the transfer function characterizing sound propagation from the lth loudspeaker and sound pressure at the point in the room $\vec{r}$, $Q^{(l)}(\omega)$ is the filter for the lth loudspeaker, and $S(\omega)$ is the input signal. For simplicity, $S(\omega)$ is assumed to be 1, and the frequency ω is omitted in what follows.

A plane wave or a sound wave generated by a monopole in free field condition is defined as the desired field in many studies because any sound fields can be expressed as the superposition of plane waves or monopoles [5–7,9–11]. In this study, the desired field is defined by a monopole:

$$\widetilde{P}_f\left(\overrightarrow{r}\right) = A\frac{e^{ikR_v}}{R_v}, \tag{2}$$

where A is the monopole amplitude, R_v is the distance between the position of the monopole $\overrightarrow{r}_v$ and a point in the control region $\overrightarrow{r}$ ($R_v = \left|\overrightarrow{r} - \overrightarrow{r}_v\right|$).
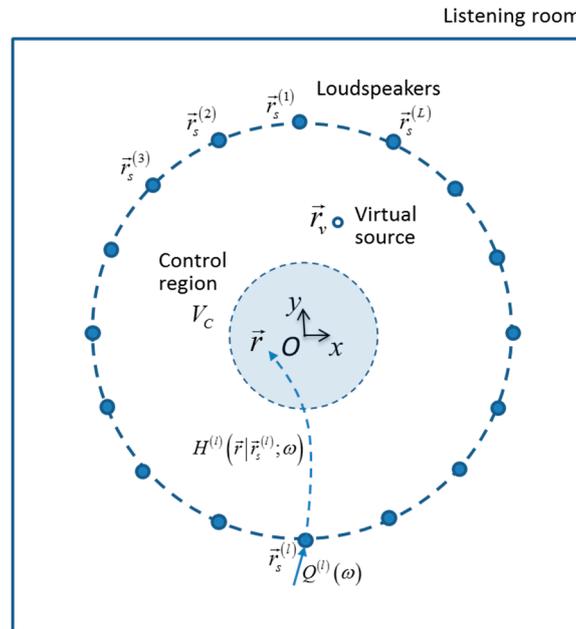

Figure 1. A typical sound field reproduction system in a listening room.

The filter $Q^{(l)}$ is obtained by the reproduction methods, such as WFS and HOA. These methods normally model the loudspeakers as monopoles or plane waves in free-field condition. Thus, the transfer function in a free field can be denoted as $H_f^{(l)}\left(\overrightarrow{r}\,\middle|\,\overrightarrow{r}_s^{(l)}\right)$. The reproduction methods attempt to make the reproduced field identical to the desired field in the control region V_c:

$$\widetilde{P}_f\left(\overrightarrow{r}\right) \approx P_f\left(\overrightarrow{r}\right) = \sum_{l=1}^{L} H_f^{(l)}\left(\overrightarrow{r}\,\middle|\,\overrightarrow{r}_s^{(l)}\right)Q^{(l)}, \quad \overrightarrow{r} \in V_c \tag{3}$$

The spatially-averaged error $\bar{e}_f$ is defined as the error between the desired and the reproduced fields:

$$\bar{e}_f^{\,2} = \frac{\int_{V_C}\left|\widetilde{P}_f\left(\overrightarrow{r}\right) - P_f\left(\overrightarrow{r}\right)\right|^2 dV}{\int_{V_C}\left|\widetilde{P}_f\left(\overrightarrow{r}\right)\right|^2 dV} \tag{4}$$

This measure quantifies how similar the reproduced field is to the desired in the space domain in terms of the pressure distribution.

2.2. Effect of Room Reflections

Loudspeaker-based reproduction systems are installed in non-anechoic conditions in most practical cases, and consequently unintended room reflections are included in the transfer function of the loudspeakers as follows,

$$H^{(l)}\left(\overrightarrow{r}\,\middle|\,\overrightarrow{r}_s^{(l)}\right) = H_f^{(l)}\left(\overrightarrow{r}\,\middle|\,\overrightarrow{r}_s^{(l)}\right) + H_{ref}^{(l)}\left(\overrightarrow{r}\,\middle|\,\overrightarrow{r}_s^{(l)}\right), \tag{5}$$

where $H_{ref}^{(l)}\left(\vec{r}\,\middle|\,\vec{r}_s^{(l)}\right)$ is the transfer function for the reverberant sound field. The reproduced field also includes the room reflections $P_{ref}\left(\vec{r}\right)$, which increase the spatial error:

$$\bar{e}^2 = \frac{\int_{V_c}\left|\widetilde{P}_f\left(\vec{r}\right) - P_f\left(\vec{r}\right) - P_{ref}\left(\vec{r}\right)\right|^2 dV}{\int_{V_c}\left|\widetilde{P}_f\left(\vec{r}\right)\right|^2 dV}. \tag{6}$$

On the contrary, room reflections are known to affect human localization of the sound source to a lesser degree according to law of the first wave front [35]. Late reflections hardly affect human localization of the sound source, and early reflections up to 10 ms can smear or shift the image of the perceived source location [35], but do not significantly change the localized direction. By removing the late reflections with a time window [41,42], the effect of reflections on the spatial error can be suppressed to some extent. Yet, some early reflections cannot be perfectly removed due to overlap, and increase the spatial error.

Hence, in non-anechoic conditions, the spatial error gives little explanations on the reproduced field [29]. This means that the spatial error cannot show whether or not there are experimental errors, such as positioning errors of microphones and loudspeakers, or difference between loudspeaker models and actual loudspeaker responses.

2.3. Effect of Amplitude Decay with Distance

The amplitude of a monopole decays with distance by $1/r$, where r is the distance to the monopole position. If a linear array or a circular array of loudspeakers is used, e.g., a 2.5D reproduction, the amplitude decay will differ. Figure 2 shows the amplitude decay in a computer simulation where the reproduced field is generated with a circular array of 16 loudspeakers at 1 kHz. The radius of the circular array is 1.5 m, and the position of a desired monopole is at (10 m, 0). The desired and the reproduced decays are compared with a special focus in the controllable zone (9.6 m < x < 10.4 m). The magnitude is normalized to have 0 dB at the center of the controllable region (x = 10). The reproduced field differs clearly from the desired decay. The difference in the amplitude decay would significantly increase the spatial error. On the contrary, the amplitude decay hardly affects the human localization in terms of perception; directional perception is mainly due to the inter-aural time difference (ITD) and the inter-aural level difference (ILD), and the distance perception is more attributed to the loudness and coloration than the amplitude decay [35–38].
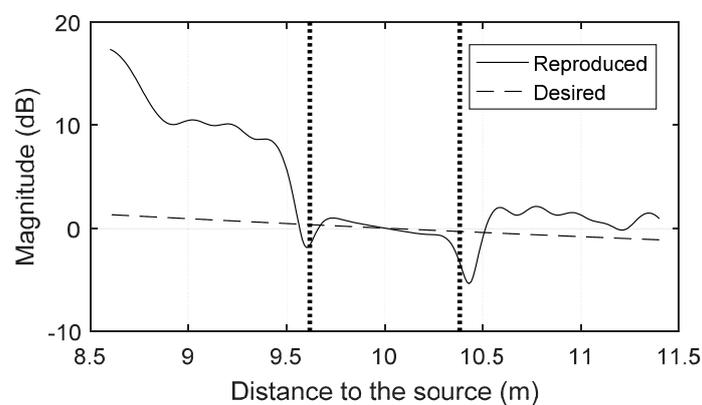

Figure 2. Amplitude decay of the reproduced and the desired field.

3. Proposed Method

3.1. Beam-Power and Directional Filter

Beamforming methods can distribute the acoustic energy in a sound field along the propagating directions. When the beamforming methods are applied to a successfully-reproduced field in a free-field condition, the beam-power should have the maximum value at the intended direction of the virtual monopole source. Even in non-anechoic conditions, the direction of the maximum value does not vary considerably because the beam-power of the direct sound is still dominant, while room reflections are distributed in many directions. The effect of the reflections can be reduced by using a directional filter that gives a greater weighting towards the monopole direction.

To obtain a consistent measure across the frequency, the main lobe needs to have a constant width. This width depends on the beamforming method and the aperture size of the sensor array. If HOA is used, the reproducible region is determined as $kr < N$ [5], where k is the wave number, N is the order of the HOA, and r is the radius of the region. This means that the aperture size decreases with the frequency. Delay-and-sum (DAS) beamforming is chosen [43], because DAS beamforming also has the narrower main lobe width at the higher frequency for a constant aperture size. Thus, HOA is used together with DAS beamforming. If the reproducible region is frequency-invariant, constant directivity beamforming methods need to be used.

The desired beamformer output with respect to the assumed azimuth angle ϕ_c for the DAS beamforming can be calculated in frequency domain as:

$$\widetilde{B}(\phi_c) = \int_0^{2\pi} \int_0^a \widetilde{P}_f(r, \phi) \exp(-ikR_c) 2\pi r \, dr \, d\phi, \tag{7}$$

where a is the radius of the controllable region of the N-th order HOA ($a = N/k$), and R_c is the distance between the measurement point (r, ϕ) and the assumed position of the source (r_c, ϕ_c), as shown in Figure 3. The intended position of the virtual monopole is known ($\vec{r}_v = (r_c, \phi_v)$) and the radius r_c is assumed to be identical to the distance of the monopole r_v. If r_v is assumed to be much greater than the radius of the reproducible region a ($r_v \gg a$) for simplicity, inserting Equation (2) into Equation (7) leads to:

$$\widetilde{B}(\phi_c) = \left| \frac{8\pi a^2 A}{r_v^2} \frac{J_1\left(2ka \sin\left(\frac{\Phi_0 - \Phi_v}{2}\right)\right)}{2ka \sin\left(\frac{\Phi_0 - \Phi_v}{2}\right)} \right|, \tag{8}$$

where J_1 is the first-order Bessel function of the first kind. See Appendix A for more details. This beam-former output is frequency-invariant as ka is constant as N. To make it independent of the distance, the normalized beam-power $\widetilde{\beta}_N$ can be defined as:

$$\widetilde{\beta}_N(\phi_c) = \frac{\widetilde{B}(\phi_c)}{\max\left\{\widetilde{B}(\phi_c)\right\}} = \left| \frac{J_1\left(2ka \sin\left(\frac{\Phi_0 - \Phi_v}{2}\right)\right)}{ka \sin\left(\frac{\Phi_0 - \Phi_v}{2}\right)} \right|, \tag{9}$$

Then, the effect of the amplitude decay with the distance vanishes.

The beam-former output of the reproduced field $B(\phi_c)$ can be obtained in the same way,

$$B(\phi_c) = \left| \int_0^{2\pi} \int_0^a P(r, \phi) \exp(-ikR_c) 2\pi r \, dr \, d\phi \right| \tag{10}$$

Unlike the ideal case (Equation (8)), the outcome of this integration is a complex number in general. The reproduced field contains the direct sound and reflections. The effect of these reflections can be reduced by a time window for excluding late reflections and a directional filter $W(\phi_c)$ for

reducing the weighting at remote directions from ϕ_v. In this study, a directional filter that has the maximum at $\phi_c = \phi_v$ is defined as:

$$W(\phi_c) = 0.5[1 + \cos(\phi_c - \phi_v)] \tag{11}$$

The beam-power error is then defined as:

$$\bar{e}_{BF}^2 = \frac{\int_0^{2\pi} \left| W(\phi_c)\beta_N(\phi_c) - W(\phi_c)\widetilde{\beta}_N(\phi_c) \right|^2 d\phi_c}{\int_0^{2\pi} \left| W(\phi_c)\widetilde{\beta}_N(\phi_c) \right|^2 d\phi_c} \tag{12}$$

where $\beta_N(\phi_c)$ is the normalized beam-power of the reproduced field.
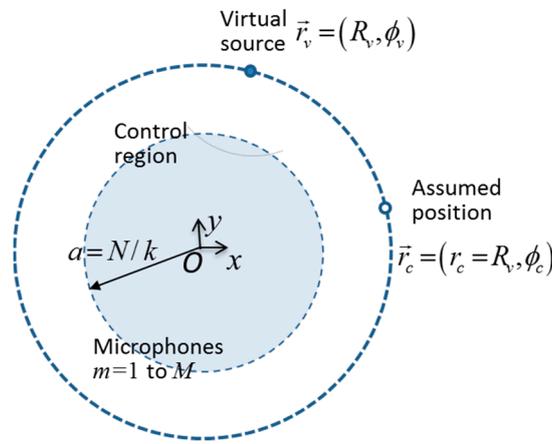
Figure 3. Setup for the beamforming method of a simple example.

3.2. Properties of the Beam-Power Error

In order to investigate the properties of the beam-power error, two sound fields are evaluated with the spatial error and the beam-power error via numerical simulations. A reference sound field has a monopole in a free field, corresponding to the desired field. To compare with the reference condition, two modifications are made: inclusion of room reflections and change in the monopole position, i.e., change in wave fronts. The position of the monopole $\vec{r}_v$ is (0, 0.9 m), and the order of HOA, N, is 7.

In the first modification, the reproduced field has room reverberation on top of the desired field. A small room (4.5 m × 4.4 m × 2.5 m) is assumed, and reflections are simulated by the image source method [44,45] with the pressure reflection coefficient, ρ, varying from 0.1 to 0.9. Late reflections that arrive after the mixing time ($t_{mix} = \sqrt{V} = 7.0$ ms) are excluded with a time window. The mixing time has been proposed as a criterion that separates early reflections from late reflections [46–48]. Figure 4 shows the beam-power distributions of the desired and the reproduced fields when the reflection coefficient is 0.9. The side lobes of the reproduced field have greater values than the desired field due to the strong reflections, but the main lobe has a small difference in these two curves, decreasing the beam-power error. Figure 5 compares the beam-power error with the spatial error with respect to the reflection coefficient. The beam-power error has at least 20 dB lower values than the spatial error, having lower values than −20 dB even for high reflection coefficients.
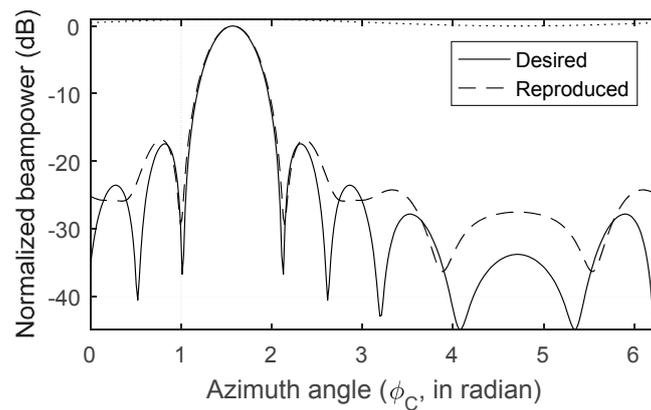
Figure 4. Normalized beam-power distributions of the desired and the reproduced fields with ρ = 0.9.
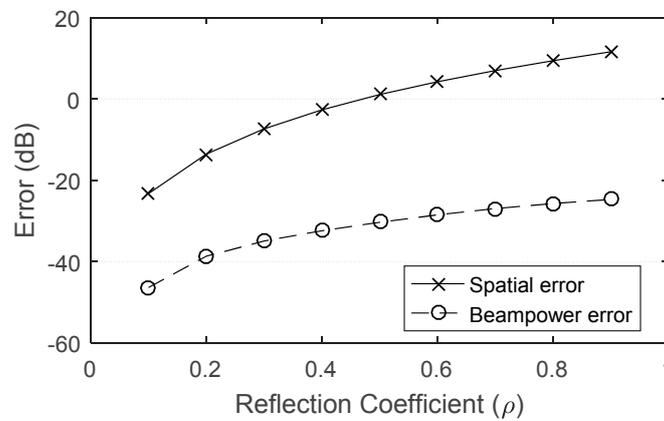

Figure 5. Spatial error and beam-power error with respect to the reflection coefficient.

In the second modification, a wrong monopole location is used for the reproduced field in an anechoic condition. This error should be effectively detected because it is induced by the error in the direction of sound waves, which can be easily perceived by human listeners. Figure 6 shows the spatial error and the beam-power error with respect to the difference of the monopole direction, $\Delta\phi$, where both errors differ by 1.7 dB on average and 2.0 dB at the most. This means that the beam-power error can detect the error due to the change in the direction of the monopole as effectively as the spatial error does.
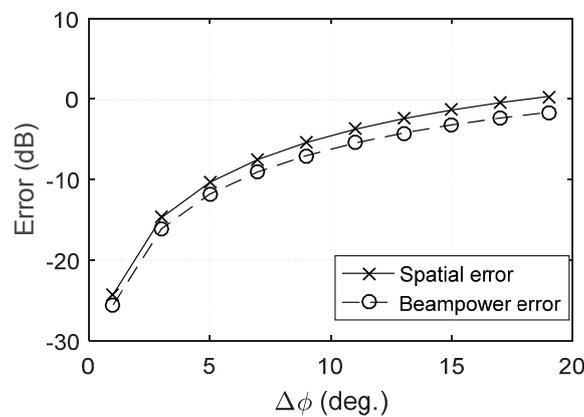

Figure 6. Spatial error and beam-power error with respect to $\Delta\phi$.

In summary, the beam-power error is much less sensitive to room reflections than the spatial error, but detects faults in the direction of arrival as effectively as the spatial error does.

4. Experimental Example

4.1. Pre-Experiment with One Source On

As a pre-experiment, the proposed measure was compared with the spatial error in a real reproduction room where only one loudspeaker was turned on. The reproduction room was an acoustically-damped room (4.5 m × 4.4 m × 2.5 m) [27,28,30] built at the Technical University of Denmark, where the reverberation time T_{30} was 0.16 s in the 125 Hz octave band, and below 0.1 s for higher frequencies. The loudspeaker was located at (0, 1.8 m). Figure 7 illustrates this setup, and the loudspeaker is indicated as #1.
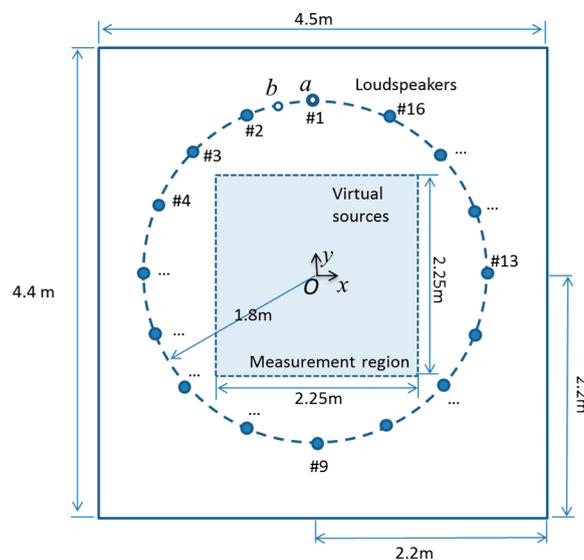

Figure 7. The listening room, loudspeaker array, and virtual sources.

The measurement was conducted with a planar array of 60 (6 × 10) microphones (B&K type 4957) at 15 positions (5 × 3). The measurement region was a square in the horizontal plane (2.25 m × 2.25 m) indicated in Figure 7, and the spacing of the measurement points was 7.5 cm. The total number of the measurement points was, thus, 900 (30 × 30) (Figure 7). The microphones were calibrated, and the phase difference among the microphones was ±5° up to 3 kHz. The microphone signals were recorded with a multi-channel analyzer (B&K frontend frame 3560D, modules 7537A and 3038B). To synchronize the measurement at each position, the signal fed into the loudspeaker was also measured. The background noise was lower than 10 dB SPL above 80 Hz, and the temperature and the humidity were 21 °C and 42%. The sampling frequency was 16,384 Hz.

The control region was assumed to be a circular region ($kr < N$) with the maximum order $N = 7$. The measurement points located in the control region were used for the beamforming and, thus, the number of the measurement points varied with the frequency. For frequencies lower than 340 Hz, the entire control region was not measured because some of the control region was outside the measurement region. The late reflections that arrived after the mixing time were excluded by the time window as used in the simulation.

Figure 8 shows the error measures with frequency. The spatial error varies from around −10 to 0 dB values at all frequencies, which is a considerable amount. Since only one loudspeaker was used, the sound field was close to that generated by a monopole at the loudspeaker position, except for room reflections, and human localization is expected to be stable because the room is acoustically

damped. Informal listening tests also confirmed a stable localization performance. In contrast, the beam-power error has values lower than –20 dB above 200 Hz. The spatial error has periodic fluctuations approximately every 280 Hz. The period of this frequency corresponds to the time delay between the direct sound and the reflection from the floor. Hence, the floor reflection is the main cause of the periodic increased spatial error, whereas the beam-power error is hardly affected by such reflections. The beam-power error is mainly affected by the difference in the angle ϕ_c of the maximum beam-power between the desired and the reproduced field.
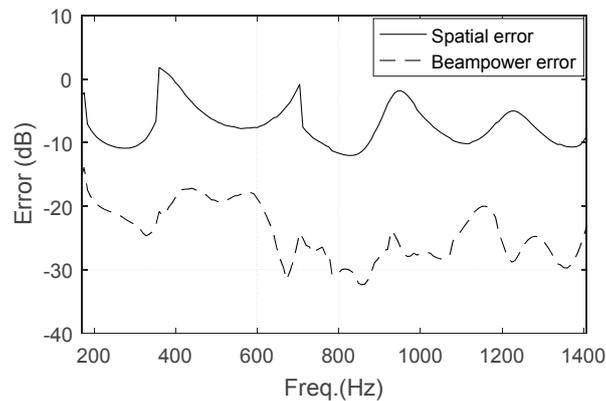

Figure 8. Spatial error and beam-power error in the pre-experiment.

4.2. Experiment with 2.5D HOA

The spatial error and the beam-power error were compared in the same room, yet with 2.5D HOA [6]. A circular array of 16 loudspeakers at equiangular positions was used, and the radius of the circle was 1.8 m (Figure 7). The loudspeakers had ± 3 dB deviation from the mean value in the entire frequency range of interest. The details of the filter used were described in Appendix B. The loudspeakers were modeled as monopoles.

Two virtual source positions were aimed for reproduction as illustrated in Figure 7. A monopole is reproduced in each experiment as indicated by '*a*' and '*b*' in Figure 7. The position '*a*' is (1.8 m, 90 degrees), and the position '*b*' is (1.8 m, 101.25 degrees). The first position corresponds to that of the first loudspeaker, and the second one is at the middle point of the first and the second loudspeakers.

Figure 9 compares the spatial error and the beam-power error for the desired monopole located at *a* (left) and *b* (right). As the position '*a*' coincide with the loudspeaker #1, one might expect a better reproduction performance than that for the position '*b*'. The spatial error, however, shows similar values and fluctuations with the frequency for those two cases. The beam-power error for position '*a*' turns out to be smaller than that for position '*b*'.
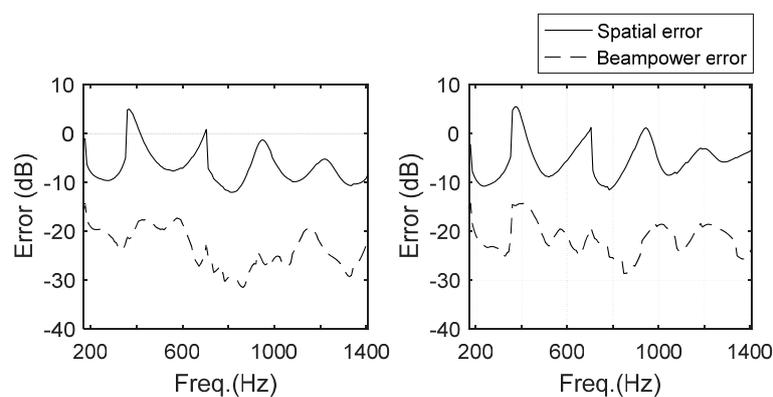

Figure 9. Spatial error and beam-power error in reproduction of a monopole at '*a*' (**left**) and '*b*' (**right**).

As explained in Figure 8, these peaks in the spatial error are ascribed to the floor reflection. On the other hand, the beam-power error is less affected by the floor reflection.

5. Discussion

A time window was used in the experiment to exclude the late reflections based on the concept of mixing time. In general, as the window becomes shorter, both the spatial and beam-power error becomes smaller because the direct sound becomes more predominant. Controlling such a window length can be useful to check whether or not the sound field is close to a free-field condition. Early reflections were included in this study, which occasionally shifted the peak in the beam-power. This might be related to the fact that early reflections of 5–10 ms can shift the perceived source position in human listening [35,49]. This relation needs to be further studied, but it is beyond the scope of the present study.

As shown in Figure 4, the beam-power distribution shows the difference between the desired and the reproduced fields with respect to the assumed direction. The shift of the propagating direction of the direct sound is shown in the main-lobe, and the effect of excessive reflections is shown mainly in the side lobes. This means that the beam-power distribution gives more information than the spatial distribution does.

In addition, as shown in Figure 6, if there are no reflections, it is expected that the beam-power error can detect the difference in the main lobe as effectively as the spatial error. Thus, for reproduced sound fields by simple panning methods, such as VBAP (vector-based amplitude panning) [50], the beam-power error is expected to be as large as the spatial error. On the other hand, because of the directional filter, the difference in the side lobes can be underestimated compared with the spatial error.

Recently, another measure called planarity has been proposed for evaluation of the reproduced fields [51]. The planarity quantifies the similarity to a plane wave sound field using the ratio between the intensity component in the direction of the plane wave and the total energy flux by the beamforming technique. Although the planarity and the proposed beam-power error use the beamforming method in common, the main difference is that the planarity does not take the shape of the main lobe into consideration because the planarity is calculated as the weighted sum of the beam-power. This makes the planarity insensitive to the difference in the wavefronts. Figure 10 shows the planarity for the case considered in Figure 6. The planarity decreases as the angle difference increases as expected, but the change is not significant. For example, the spatial error and the beam-power error increases by around 15 dB for the shift of five degrees, while the planarity decreases from 0.970 to 0.967. Consequently, it is concluded that the beam-power error is more effective to detect the difference in wave fronts.
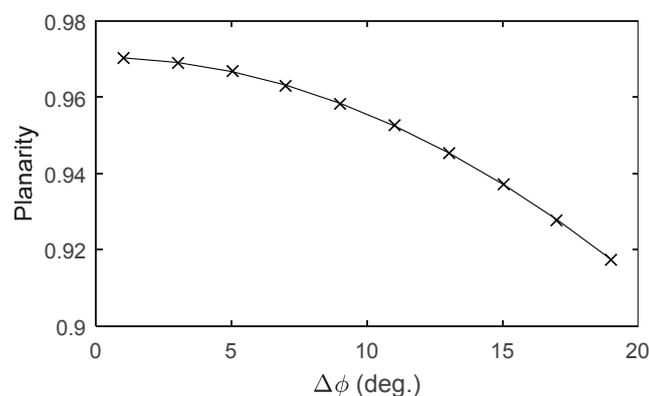

Figure 10. Planarity with respect to the $\Delta \phi$ in the same condition of Figure 6.

The experimental example used a specific reproduction method (2.5D HOA) in a specific listening room, and sound fields in the controllable region on the horizontal plane, which is frequency dependent ($kr < N$), were used to obtain the beam-power distribution. However, this measure can be used in other

circumstances. For example, a frequency-independent spherical region that just includes the listener's head can be chosen, because the sound field in this region is directly related to the sound pressure that the listener would have at two ears. If 3D reproduction methods are used, and the elevation angle of the virtual source is also of concern, the beamformer output can be extended to the two-dimensional case that the beamformer output is expressed as $B(\theta_c, \phi_c)$.

The present study is limited to the physical evaluation of sound fields. Subjective tests are needed to prove that the proposed measure has good agreements with human localization in future works.

6. Conclusions

The present study proposed a measure that can quantify how well a sound field is reproduced based on a beamforming technique. Instead of directly comparing the reproduced field pressure with that of the desired field, as the spatial error does, the proposed measure compares the beam-power between the reproduced and desired sound field. The spatial error is overly affected by the amplitude decay with distance and early reflections in rooms and overestimated for a simple sound field generated by one loudspeaker in a highly-damped room. The proposed beam-power measure is less sensitive to the amplitude decay over distance and room reflections, which could better correspond with human perception. The beam-power error can pick up the difference in wavefronts of the direct sound in an equally effective manner as the spatial error. This measure can be useful particularly when a reproduction system is installed in non-anechoic rooms or a 2.5D reproduction is conducted with a linear or a circular array of loudspeakers.

Author Contributions: Ji-Ho Chang designed, performed the experiments, and analyzed the data; Ji-Ho Chang and Cheol-Ho wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A Beamforming Output of a Circular Region

The distance R_c in Equation (2) can be approximated as:

$$R_c = \sqrt{r_s{}^2 + r^2 - 2r_s r \cos(\phi_c - \phi)}$$
$$\cong r_s - r \cos(\phi_s - \phi), \quad if \; r_s \gg a. \tag{A1}$$

The desired field in free field $\widetilde{P}_f(r, \phi)$ can also be approximated as:

$$\widetilde{P}_f(r, \phi) = A \frac{\exp(ikR_c)}{R_c}$$
$$\cong A \frac{\exp(ikr_s)}{r_s} \exp(-ikr \cos(\phi_0 - \phi)). \tag{A2}$$

Inserting Equations (A1) and (A2) into Equation (7):

$$\widetilde{B}(\phi_c) = \frac{A}{r_s{}^2} \int_0^{2\pi} \int_0^a \exp(ikr[\cos(\phi_c - \phi) - \cos(\phi_0 - \phi)]) 2\pi r dr d\phi$$
$$= \frac{A}{r_s{}^2} \int_0^{2\pi} \int_0^a \exp\left(ikr\left[2\sin\left(\frac{\phi_c + \phi_0}{2} - \phi\right)\sin\left(\frac{\phi_0 - \phi_c}{2}\right)\right]\right) 2\pi r dr d\phi. \tag{A3}$$

ϕ can be substituted by $\phi' = \frac{\pi}{2} - \frac{\phi_s + \phi_0}{2} + \phi$:

$$\widetilde{B}(\phi_c) = \frac{A}{r_s{}^2} \int_0^a \int_0^{2\pi} \exp\left(i2kr \sin\left(\frac{\phi_0 - \phi_c}{2}\right) \cos \phi'\right) d\phi' 2\pi r dr$$
$$= \frac{8\pi A}{r_s{}^2} \int_0^a \int_0^{\pi/2} \cos\left(2kr \sin\left(\frac{\phi_0 - \phi_c}{2}\right) \cos \phi'\right) d\phi' r dr. \tag{A4}$$

Using the properties of the Bessel function J_n, this equation is reduced as:

$$
\begin{aligned}
\widetilde{B}(\phi_c) &= \frac{8\pi A}{r_s^2} \int_0^a J_0\left(2kr\sin\left(\frac{\Phi_0-\Phi_s}{2}\right)\right) r\,dr \\
&= \frac{8\pi A}{r_s^2} \left[\frac{r^2 J_1\left(2kr\sin\left(\frac{\Phi_0-\Phi_s}{2}\right)\right)}{2kr\sin\left(\frac{\Phi_0-\Phi_s}{2}\right)}\right]_0^a \\
&= \frac{8\pi a^2 A}{r_s^2} \frac{J_1\left(2ka\sin\left(\frac{\Phi_0-\Phi_s}{2}\right)\right)}{2ka\sin\left(\frac{\Phi_0-\Phi_s}{2}\right)}
\end{aligned}
\tag{A5}
$$

Appendix B Near-Field Compensated Higher-Order Ambisonics

The reproduction method, 2.5D NFC-HOA, can generate a monopole source with a circular array of loudspeakers [30]. The distance and the angle of the monopole source can be controlled in the horizontal plane of the loudspeakers. The loudspeakers are modeled as monopoles.

The NFC-HOA filter can be written with complex cylindrical harmonics as follows. For each monopole source located at (R_v, ϕ_v) (in polar coordinates). Ambisonic signals were encoded up to an order N and decoded to an array of L loudspeakers located at $(r_s, \phi_s^{(l)})$. The filter for the lth loudspeaker $Q_{NFC}^{(l)}$ can be written as:

$$
Q_{NFC}^{(l)} = \sum_{n=-N}^{N} W_n H_n^{NFC(R_s,\phi_s)} e^{in\phi_v} e^{-in\phi_s^{(l)}}.
\tag{A6}
$$

$H_n^{NFC(R_v,\phi_v)}$ is the NFC filter [6]:

$$
H_n^{NFC(R_s,\phi_s)} = \frac{h_n^{(1)}(kR_v)}{h_0^{(1)}(kR_v)} \frac{h_0^{(1)}(kr_s)}{h_n^{(1)}(kr_s)},
\tag{A7}
$$

where $h_n^{(1)}$ is the spherical Hankel function of the first kind, and W_n is the regularization function [26]:

$$
W_n = \frac{2}{\left|H_n^{NFC(R_v,\phi_v)}\right|^2 + 1}.
\tag{A8}
$$

The number of the coefficients for a given N in Equation (A6) is $2N + 1$. It is well known that the number of loudspeakers should be greater than $2N + 1$ for an accurate reproduction [5]. In this experiment, 16 loudspeakers were used, which allow for a maximum order of $N = 7$.

References

1.   Fellgett, P.B. Ambisonic reproduction of directionality in surround-sound systems. *Nature* **1974**, *252*, 534–538. [CrossRef]
2.   Gerzon, M.A. Periphony: Width-height sound field reproduction. *J. Audio Eng. Soc.* **1973**, *21*, 2–10.
3.   Gerzon, M.A. Ambisonics in multichannel broadcasting and video. *J. Audio Eng. Soc.* **1985**, *33*, 859–871.
4.   Bamford, J.S.; Vanderkooy, J. Ambisonic sound for us. In Proceedings of the 99th Audio Engineering Society Convention, New York, NY, USA, 6–9 October 1995.
5.   Ward, D.B.; Abhayapala, T.D. Reproduction of a plane-wave sound field using an array of loudspeakers. *IEEE Trans. Speech Audio Proc.* **2001**, *9*, 697–707. [CrossRef]
6.   Daniel, J. Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format. In Proceedings of the 23th Audio Engineering Society International Conference, Helsingør, Denmark, 23–25 May 2003.
7.   Poletti, M.A. Three-dimensional surround sound systems based on spherical harmonics. *J. Audio Eng. Soc.* **2005**, *53*, 1004–1025.

8. Wu, Y.J.; Abhayapala, T.D. Theory and design of soundfield reproduction using continuous loudspeaker concept. *IEEE Trans. Audio Speech Lang. Proc.* **2009**, *17*, 107–116. [CrossRef]

9. Berkhout, A. A holographic approach to acoustic control. *J. Audio Eng. Soc.* **1988**, *36*, 977–995.

10. Berkhout, A.; de Vries, D.; Vogel, P. Acoustic control by wave field synthesis. *J. Acoust. Soc. Am.* **1993**, *93*, 2764–2778. [CrossRef]

11. Spors, S.; Rabenstein, R.; Ahrens, J. The theory of wave field synthesis revisited. In Proceedings of the 124th Convention Audio Engineering Society, Amsterdam, The Netherlands, 17–20 May 2008.

12. Wu, Y.J.; Abhayapala, T.D. Spatial multizone soundfield reproduction: Theory and design. *IEEE Trans. Audio Speech Lang. Process.* **2011**, *19*, 1711–1720. [CrossRef]

13. Poletti, M.A.; Fazi, F.M. An approach to generating two zones of silence with application to personal sound systems. *J. Acoust. Soc. Am.* **2015**, *137*, 598–605. [CrossRef] [PubMed]

14. Jin, W.; Kleijn, W.B. Theory and design of multizone soundfield reproduction using sparse methods. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2015**, *23*, 2343–2355.

15. Zhang, W.; Abhayapala, T.D.; Betlehem, T.; Fazi, F.M. Analysis and control of multi-zone sound field reproduction using modal-domain approach. *J. Acoust. Soc. Am.* **2016**, *140*, 2134–2144. [CrossRef] [PubMed]

16. De Vries, D. *Wave Field Synthesis*; Audio Engineering Society Inc.: New York, NY, USA, 2009; pp. 44–83.

17. De Vries, D.; Vogel, P. Experience with a sound enhancement system based on wave field synthesis. In Proceedings of the 95th Convention Audio Engineering Society, New York, NY, USA, 7–10 October 1993.

18. Corteel, E.; Nguyen, K.-V.; Warusfel, O.; Caulkins, T.; Pellegrini, R. Objective and subjective comparison of electrodynamic and MAP loudspeakers for wave field synthesis. In Proceedings of the 30th Conference Audio Engineering Society, Saariselka, Finland, 15–17 March 2007.

19. Klehs, B.; Sporer, T. Wave field synthesis in the real world: Part 1—In the living room. In Proceedings of the 114th Convention Audio Engineering Society, Amsterdam, The Netherlands, 1 March 2003.

20. Sporer, T.; Klehs, B. Wave field synthesis in the real world: Part 2—In the movie theatre. In Proceedings of the 116th Convention Audio Engineering Society, Berlin, Germany, 8–11 May 2004.

21. Buchner, H.; Spors, S.; Kellermann, W. Full-Duplex systems for sound field recording and Auralization based on wave field synthesis. In Proceedings of the 116th Convention Audio Engineering Society, Berlin, Germany, 1 May 2004.

22. Geier, M.; Wierstorf, H.; Ahrens, J.; Wechsung, I.; Raake, A.; Spors, S. Perceptual evaluation of focused sources in wave field synthesis. In Proceedings of the 128th Convention Audio Engineering Society, London, UK, 1 May 2010.

23. Lopez, J.J.; Cobos, M.; Pueo, B.; Aguilera, E. Wave field synthesis for next generation videoconferencing. In Proceedings of the 4th International Symposium Communications, Control and Signal Processing, (ISCCSP 2010), Limassol, Cyprus, 3–5 March 2010; pp. 1–4.

24. Farina, A.; Capra, A.; Martignon, P.; Fontana, S.; Adriaensen, F.; Galaverna, P.; Malham, D. Three-dimensional acoustic displays in a museum employing WFS (wave field synthesis) and HOA (high order Ambisonics). In Proceedings of the 14th International Congress on Sound and Vibration (ICSV 14), Cairns, Australia, 9–12 July 2007.

25. Fazi, F.M.; Nelson, P.A.; Christensen, J.E.N.; Seo, J. Surround system based on three-dimensional sound field reconstruction. In Proceedings of the 125th Convention Audio Engineering Society, San Francisco, CA, USA, 1 October 2008.

26. Epain, N.; Guillon, P.; Kan, A.; Kosobrodov, R.; Sun, D.; Jin, C.; van Schaik, A. Objective evaluation of a three-dimensional sound field reproduction system. In Proceedings of the 20th International Congress on Acoustics (ICA 2010), Sydney, Australia, 23–27 August 2010.

27. Favrot, S.; Buchholz, J.M. LoRA: A loudspeaker-based room Auralization system. *Acta Acust. United Acust.* **2010**, *96*, 364–375. [CrossRef]

28. Kaesbach, J.; Favrot, S.; Buchholz, J. Evaluation of a Mixed-Order Planar and Periphonic Ambisonics Playback Implementation. Available online: http://johannes.kaesbach.de/Acoustics_files/PaperMOA_EAAtemplate2.pdf (accessed on 3 March 2017).

29. Solvang, A. Spectral impairment of two-dimensional higher order ambisonics. *J. Audio Eng. Soc.* **2008**, *56*, 267–279.

30. Favrot, S.; Buchholz, J. Reproduction of nearby sound sources using higher-order ambisonics with practical loudspeaker arrays. *Acta Acust. United Acust.* **2012**, *98*, 48–60. [CrossRef]

31. Guastavino, C.; Katz, B.F.G. Perceptual evaluation of multi-dimensional spatial audio reproduction. *J. Acoust. Soc. Am.* **2004**, *116*, 1105–1115. [CrossRef] [PubMed]

32. Frank, M.; Zotter, F.; Sontacchi, A. Localization experiments using different 2D ambisonics decoders. In Proceedings of the 25th Tonmeistertagung, Leipzig, Germany, 13–16 November 2008.

33. Bertet, S.; Daniel, J.; Parizet, E.; Gros, L.; Warusfel, O. Investigation of the perceived spatial resolution of higher order ambisonic sound fields: A subjective evaluation involving virtual and real 3D microphones. In Proceedings of the 30th Convention Audio Engineering Society, Saariselka, Finland, 15–17 March 2007.

34. Chang, J.H.; Choi, J.W.; Kim, Y.H. A plane wave generation method by wave number domain point focusing. *J. Acoust. Soc. Am.* **2010**, *128*, 2758–2767. [CrossRef] [PubMed]

35. Blauert, J. *Spatial Hearing: The Psychophysics of Human Sound Localization*; MIT Press: Cambridge, MA, USA, 1983; pp. 222–235.

36. Gerzon, M.A. The design of distance panpots. In Proceedings of the 92nd Convention Audio Engineering Society, Vienna, Austria, 1 March 1992.

37. Shinn-Cunningham, B.G.; Kopco, N.; Martin, T.J. Localizing nearby sound sources in a classroom: Binaural room impulse responses. *J. Acoust. Soc. Am.* **2005**, *117*, 3100–3115. [CrossRef] [PubMed]

38. Pellegrini, R.S.; Horbach, U. Perception-based design of virtual rooms for sound reproduction. In Proceedings of the 112nd Convention Audio Engineering Society, Munich, Germany, 10–13 May 2002.

39. Betlehem, T.; Abhayapala, T.D. Theory and design of sound field reproduction in reverberant rooms. *J. Acoust. Soc. Am.* **2005**, *117*, 2100–2111. [CrossRef] [PubMed]

40. Jin, W.; Kleijn, W.B. Multizone sound field reproduction in reverberant rooms using compressed sensing techniques. In Proceedings of the IEEE ICASSP, Florence, Italy, 4–9 May 2014; pp. 4728–4732.

41. Farina, A. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In Proceedings of the 108th Convention Audio Engineering Society, Paris, France, 19–22 February 2000.

42. Muller, S.; Massarani, P. Transfer-function measurement with sweeps. *J. Audio Eng. Soc.* **2001**, *49*, 443–471.

43. Dudgeon, D.; Johnson, D. *Array Signal Processing: Concepts and Techniques*; Prentice Hall: Englewood Cliffs, NJ, USA, 1993; pp. 111–119.

44. Allen, J.B.; Berkley, D.A. Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Am.* **1979**, *65*, 943–950. [CrossRef]

45. Lee, H.; Lee, B.-H. An efficient algorithm for the image model technique. *Appl. Acoust.* **1988**, *24*, 87–115. [CrossRef]

46. Polack, J.-D. Playing billiards in the concert hall: The mathematical foundations of geometrical room acoustics. *Appl. Acoust.* **1993**, *38*, 235–244. [CrossRef]

47. Defrance, G.; Polack, J.-D. Estimating the mixing time of concert halls using the eXtensible Fourier Transform. *Appl. Acoust.* **2010**, *71*, 777–792. [CrossRef]

48. Jeong, C.-H.; Brunskog, J.; Jacobsen, F. Room acoustic transition time based on reflection overlap (L). *J. Acoust. Soc. Am.* **2010**, *127*, 2733–2736. [CrossRef] [PubMed]

49. Barron, M.F.E. *Auditorium Acoustics and Architectural Design*; E & FN Spon: London, UK, 1993; Chapter 3.

50. Pulkki, V. Virtual sound source positioning using vector base amplitude panning. *J. Audio Eng. Soc.* **1997**, *45*, 456–466.

51. Jackson, P.J.; Jacobsen, F.; Coleman, P.; Pedersen, J.A. Sound field planarity characterized by superdirective beamforming. In Proceedings of the International Congress on Acoustics, Montreal, QC, Canada, 2–7 June 2013.