

Perceptual analysis of directional late reverberation

Benoit Alary, Pierre Massé, Sebastian J Schlecht, Markus Noisternig, Vesa Välimäki

▶ **To cite this version:**

Benoit Alary, Pierre Massé, Sebastian J Schlecht, Markus Noisternig, Vesa Välimäki. Perceptual analysis of directional late reverberation. Journal of the Acoustical Society of America, 2021, 149 (5), pp.3189 - 3199. 10.1121/10.0004770 . hal-03251692

HAL Id: hal-03251692

https://hal.sorbonne-universite.fr/hal-03251692

Submitted on 7 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Perceptual analysis of directional late reverberation

Benoit Alary, Pierre Massé, Sebastian J. Schlecht, Markus Noisternig, and Vesa Välimäki

Citation: The Journal of the Acoustical Society of America **149**, 3189 (2021); doi: 10.1121/10.0004770
View online: https://doi.org/10.1121/10.0004770
View Table of Contents: https://asa.scitation.org/toc/jas/149/5
Published by the Acoustical Society of America

ARTICLES YOU MAY BE INTERESTED IN

Perceptual implications of different Ambisonics-based methods for binaural reverberation
The Journal of the Acoustical Society of America **149**, 895 (2021); https://doi.org/10.1121/10.0003437

Seminal article about model-based space-time array processing
The Journal of the Acoustical Society of America **149**, R9 (2021); https://doi.org/10.1121/10.0004816

Machine learning in acoustics: Theory and applications
The Journal of the Acoustical Society of America **146**, 3590 (2019); https://doi.org/10.1121/1.5133944

Acoustics Apps: Interactive simulations for digital teaching and learning of acoustics
The Journal of the Acoustical Society of America **149**, 1175 (2021); https://doi.org/10.1121/10.0003438

The role of early and late reflections on perception of source orientation
The Journal of the Acoustical Society of America **149**, 2255 (2021); https://doi.org/10.1121/10.0003823

Sensitivity analysis of pinna morphology on head-related transfer functions simulated via a parametric pinna model
The Journal of the Acoustical Society of America **149**, 2559 (2021); https://doi.org/10.1121/10.0004128

ASA ACOUSTICAL SOCIETY OF AMERICA
Across Acoustics
The official podcast highlighting authors' research from our publications

JASA ARTICLE

Perceptual analysis of directional late reverberation

Benoit Alary,[1,a)] Pierre Massé,[2,b)] Sebastian J. Schlecht,[3,c)] Markus Noisternig,[2,d)] and Vesa Välimäki[1,e)]
[1]*Department of Signal Processing and Acoustics, Aalto University, FI-02150 Espoo, Finland*
[2]*Sciences et Technologies de la Musique et du Son (STMS)–IRCAM, CNRS, Sorbonne Université, 75004 Paris, France*
[3]*Department of Signal Processing and Acoustics and Department of Media, Aalto University, FI-02150 Espoo, Finland*

ABSTRACT:
The late reverberation characteristics of a sound field are often assumed to be perceptually isotropic, meaning that the decay of energy is perceived as equivalent in every direction. In this paper, we employ Ambisonics reproduction methods to reassess how a decaying sound field is analyzed and characterized and our capacity to hear directional characteristics within late reverberation. We propose the use of objective measures to assess the anisotropy characteristics of a decaying sound field. The energy-decay deviation is defined as the difference of the direction-dependent decay from the average decay. A perceptual study demonstrates a positive link between the range of these energy deviations and their audibility. These results suggest that accurate sound reproduction should account for directional properties throughout the decay.
© *2021 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).* https://doi.org/10.1121/10.0004770
(Received 22 September 2020; revised 31 March 2021; accepted 2 April 2021; published online 12 May 2021)
[Editor: Ning Xiang] Pages: 3189–3199

I. INTRODUCTION

Artificial reverberation aims to reproduce the perceptual effect of sound propagating in a room (Schroeder and Logan, 1961; Välimäki *et al.*, 2012). In multichannel sound reproduction, the late reverberation is usually simplified to a set of decorrelated signals approximating a diffuse field (Gerzon, 1972; Schroeder and Logan, 1961). A diffuse field is a theoretical state that occurs after the reflection of sound in a space creates a near infinite number of incoherent plane waves that are evenly spread out. These plane waves create a diffuse distribution of energy, which means it is statistically equivalent at all locations and in every direction. These two properties are known, respectively, as homogeneity and isotropy (Kuttruff, 2009).

However, these idealized conditions never truly exist in practice. In a real room, the shapes, materials, and surfaces it contains all influence the diffusion of energy (Balachandran and Robinson, 1967; Kuttruff, 2009). Anisotropy, or an uneven distribution of energy in different directions, is the main challenge in the design of a reverberation chamber (Balachandran and Robinson, 1967; D'Antonio *et al.*, 2018; Nolan *et al.*, 2018; Nolan *et al.*, 2020; Pierce, 1974). This paper studies the perceptual implications of anisotropy in the reproduction of late reverberation.

Diffuseness is a measure that estimates to what extent the conditions of an ideal diffuse field are satisfied. However, diffuseness is not a uniquely defined measure, as there are different ways to quantify it (Epain and Jin, 2016). Some definitions look at the acoustic energy (Gover *et al.*, 2002, 2004), whereas others rely on the acoustic intensity (Kuttruff, 2009; Pulkki, 2007) or the covariance in the spherical harmonic domain (SHD) (Epain and Jin, 2016). In this study, we use a diffuseness measure that makes no assumptions about the isotropy of the late reverberation. Specifically, we consider a normalized spatial coherence measure inspired by Epain and Jin (2016), which is adapted to the context of a spatial decomposition, as described in Massé *et al.* (2020a).

The mixing time, which is an important descriptor in artificial reverberation, refers to the transition point between specular and diffuse conditions and can be estimated from an impulse response (Jot *et al.*, 1997; Polack, 1993; Schlecht and Habets, 2017). In Götz *et al.* (2015), the mixing time is estimated based on the diffuseness, whereas in Lindau *et al.* (2012), a listening experiment compares the perceived mixing time to existing estimation methods in a binaural reproduction system. After the mixing time, a decaying sound field is typically considered perceptually isotropic (Lindau *et al.*, 2012; Polack, 1993; Schlecht and Habets, 2017; Blesser, 2001).

In accordance with this assumption, Schroeder suggested that the main requirement for multichannel artificial reverberators was to produce a set of low-correlated signals (Schroeder, 1962; Schroeder and Logan, 1961). These were obtained by combining the signals from different delay paths. This principle was carried over as more sophisticated delay networks were introduced to formalize multichannel reverberation, where different signal paths from the same

[a)]Electronic mail: Benoit.Alary@aalto.fi, ORCID: 0000-0002-3825-4941.
[b)]ORCID: 0000-0001-8083-6358.
[c)]ORCID: 0000-0001-8858-4642.
[d)]ORCID: 0000-0003-1347-9826.
[e)]ORCID: 0000-0002-7869-292X.

system can be used to obtain decorrelated signals and to create isotropic decays (Gerzon, 1972; Jot and Chaigne, 1991; Välimäki *et al.*, 2012). Similarly, Xiang *et al.* (2019) suggest using pseudo-random signals and spectral envelopes to control the correlation between the two channels of a binaural reverberation algorithm. Only in recent work were delay networks used to produce anisotropic decay characteristics (Alary and Politis, 2020; Alary *et al.*, 2019b).

In Lachenmayr *et al.* (2016), a perceptual study confirmed that spatial features of the late reverberation contribute to the feeling of envelopment of a listener, whereas Romblom *et al.* (2016) demonstrated our capacity to hear direction-dependent variations in a sound field. Luizard *et al.* (2015) showed the perceptual threshold for double-slope reverberation present in coupled spaces. Objective analysis methods have also been developed to analyze the direction-dependent characteristics of a decaying sound field (Alary *et al.*, 2019a; Berzborn and Vorländer, 2018; Nolan *et al.*, 2018; Sakuma and Eda, 2013). To inform the development of spatial audio algorithms and the necessity of reproducing directional late reverberation, assessing the perceptual threshold of these characteristics is essential.

This paper proposes an analysis method that extracts direction-dependent energy characteristics from a spatial impulse response (SIR) encoded in the SHD. The energy-decay deviation (EDD) is proposed as a measure to calculate direction-dependent deviations in the energy decay. Through this approach, the EDD can highlight the anisotropic features of a SIR. A subjective evaluation method capable of assessing our capacity to detect changes in the directivity of late reverberation is also detailed. Using this method, a perceptual study is conducted, and the connection between the analysis method and the subjective results is discussed.

This paper is organized as follows: Sec. II covers background information relevant to the analysis method and the perceptual study. Section III introduces the EDD as an objective method to analyze anisotropic decay and shows example results. In Sec. IV, a subjective evaluation method is proposed and performed with a selected set of SIRs, along with an analysis of its results. Finally, Sec. V discusses future work and research directions and concludes the paper.

II. BACKGROUND

A. Mixing-time estimation

The mixing time t_{mix} specifies the moment where an impulse response transitions from early reflections to late reverberation (Jot *et al.*, 1997; Polack, 1993). We can exploit a measure of coherence to estimate the mixing time for SIRs (Götz *et al.*, 2015). Since we must consider coherence uncoupled from assumptions of isotropy, measures of diffuseness calculated directly in the SHD, such as the pseudo-intensity vector measure (Ahonen and Pulkki, 2009), the signal-to-diffuse ratio (Jarrett *et al.*, 2012), or COMEDIE (Epain and Jin, 2016), are inapplicable.

However, an analog of the COMEDIE measure, which is based on an eigendecomposition of the covariance matrix in the SHD, can be defined in the spatial domain. To dissociate coherence from the spatial power distribution, a normalized covariance must be calculated for each frequency (Massé *et al.*, 2020a). The results have values ranging between 0 and 1, with 0 corresponding to a fully coherent sound field and 1 to a perfectly incoherent one. Plotting the temporal evolution of this measure leads to an incoherence profile (Fig. 1).

Incoherence profiles generally begin with low values, due to the specular early reflections, and rapidly increase before reaching a relatively stable maximum value (Fig. 1). One way to define t_{mix} is as the moment this stable maximum is reached and the incoherence profile shows no more interference from discrete reflections, provided that the maximum is sufficiently high (e.g., ≥ 0.75) (Massé *et al.*, 2020a).

To identify the moment this stable maximum is reached, an adaptive Ramer–Douglas–Peucker algorithm can be used to segment the incoherence profile (Prasad *et al.*, 2012). Developed for dominant-point detection in digital image-processing, the algorithm aims to fit an arbitrary curve using linear segments. The segmentation is determined through recursive linear regressions and an adaptive maximum deviation threshold. Here, this enables the identification of the aforementioned sections, i.e., the arrival of early reflections, during which time the incoherence measure quickly increases, and the onset of the stable maximum value maintained throughout the late reverberation tail (Massé *et al.*, 2020a). In practice, this maximum value is not exactly constant due to the increasing presence of background noise; as such, the noise-floor time (see Sec. II C) must be estimated first to avoid erroneously identifying the noise floor instead of the mixing time.

This first segmentation is shown with circles in Fig. 1. The segment with the smallest slope and covering the

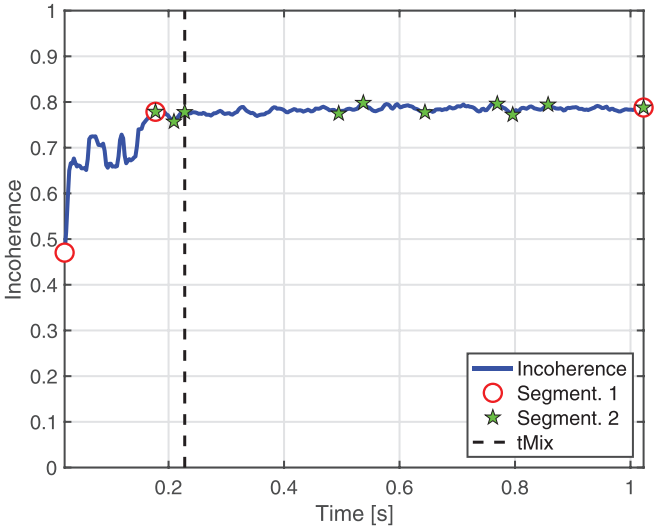


FIG. 1. (Color online) example of mixing-time estimation using a SIR of the Staatstheater, as detailed in Sec. III C 4.

https://doi.org/10.1121/10.0004770

longest duration corresponds to the maximum incoherent segment and thus the late reverberation. The mixing time is then tentatively defined as the start of this segment.

The selected late reverberation segment of the incoherence profile can then itself be re-segmented to detect any irregularities near its onset (green stars in Fig. 1), which may correspond to late-arriving coherent early reflections overlooked by the original segmentation (Massé *et al.*, 2020a). The mixing time may then be re-adjusted accordingly by giving the resulting sub-segments "scores" calculated as the geometric mean of each sub-segment's length and the inverse of its slope's absolute value. Choosing to re-adjust the mixing time to the start of the first sub-segment whose score is above or equal to the median score has been found to give consistently robust results.

B. Energy-decay curve (EDC)

The EDC was first introduced as a way to calculate the decay time T_{60} from a room impulse response (Schroeder, 1965). The EDC consists of the reverse integration of the energy from an impulse response h, which can be calculated at time t using

$$\mathrm{EDC}(t) = \int_t^\infty h^2(\tau)\,\mathrm{d}\tau. \tag{1}$$

The EDC was later expanded to the time-frequency domain with the energy-decay relief (EDR) (Jot, 1992), in which a set of frequency bands $\boldsymbol{\omega}$ are used to calculate frequency-dependent decay curves using

$$\mathrm{EDR}(t,\omega) = \int_t^\infty h^2(\tau,\omega)\,\mathrm{d}\tau. \tag{2}$$

C. Noise-floor time

One weakness of the EDC as an analysis method is the contribution to the reverse integration of non-decaying background noise present in a measured impulse response. As such, a long period of noise contained in an impulse response will have an impact on the EDC calculation and hide some details from the decay analysis (Guski and Vorländer, 2014; Karjalainen *et al.*, 2002). Therefore, when analyzing an impulse response, identifying the moment where this noise becomes prominent in relation to the decaying signal is important.

One method to estimate the noise-floor time t_{noise} is to analyze the EDC on the dB scale using an adaptive Ramer–Douglas–Peucker algorithm to yield simplified curve segments (Massé *et al.*, 2020b). In the case of a SIR, the omnidirectional channel can be used. These segments are then compared to the reverse integration of an ideal non-decaying dB-scale noise profile to find the best-matching one. Sufficient headroom above the noise profile should be used to account for the transition period where the reverberation decay and the noise floor start to blend. The noise-floor limit t_{noise} is defined as the last segment point above this specified headroom.

D. Denoising

Noise can also create undesired artefacts when using an impulse response for sound reproduction through convolution, since audible noise will create an infinite reverberation effect. Furthermore, amplifying the impulse response will also amplify the noise in the reproduction, which limits its usable dynamic range. For these reasons, methods have been developed to replace the end of an impulse response with an artificial decaying noise sequence, which follows the appropriate decay (Massé *et al.*, 2020b).

In the case of isotropic diffuse reverberation tails, this denoising can be implemented directly in the SHD (Massé *et al.*, 2020b). However, to allow for anisotropic decays, a spatial decomposition must be used, such as a plane wave decomposition (Massé *et al.*, 2020a). This decomposition must be designed to preserve linear independence between the signals, maximize spatial incoherence (see Sec. II A), and minimize directivity variance over the sphere (Massé *et al.*, 2020a). The first condition implies that the number of decomposition directions must be exactly equal to the number of SHD components in order to transform back to the SHD after denoising. The second and third conditions have been found to be jointly optimized by using a Fliege-type layout for the decomposition directions (Fliege and Maier, 1996).

III. OBJECTIVE EVALUATION

To gain a better understanding of the direction-dependent characteristics of a SIR, we propose an objective measurement method to analyze the energy distribution in a sound field throughout its decay.

A. Proposed method

Starting from a SIR encoded in Ambisonics, which is a compact representation of the sound field in the SHD (Gerzon, 1975), we first extract a set of directional impulse responses (DIRs) for a chosen set of incident directions. A DIR is obtained from the SIR using a beamforming method. Hypercardioid beamforming is used here for its simplicity and because it can extract a signal with a maximum directivity index (Rafaely, 2015). A signal obtained with the hypercardioid beamformer can be formulated as

$$y(t,\phi,\theta) = \mathbf{y}(\phi,\theta)\,\mathbf{s}(t), \tag{3}$$

where $\mathbf{y}(\phi,\theta)$ is the $(L+1)$-dimensional vector of spherical harmonic functions Y_{lm} of order $l \in \mathbb{Z}^+$ and degree $m \in [-l,l]$, up to a band limit $l \le L$ for a given direction with longitudinal value $0 \le \phi < 2\pi$ and elevation value $-\pi/2 \le \theta \le \pi/2$, and $\mathbf{s}(t)$ is the Ambisonics signal at time t.

The beamformer used to extract the DIRs will impact the data used in the analysis. More specifically, the width of the mainlobe will have a smoothing effect over multiple

directions and will reduce the dynamic range of individual DIR signals when a narrow characteristic is present in a particular direction.

From these DIRs, we can calculate directional EDCs (Berzborn and Vorländer, 2018) by updating Eq. (1) to

$$\mathrm{EDC}(t,\phi,\theta) = \int_t^\infty y^2(\tau,\phi,\theta)\,\mathrm{d}\tau. \tag{4}$$

These energy curves are converted to the dB scale:

$$\mathrm{EDC}_{\mathrm{dB}}(t,\phi,\theta) = 10\,\log_{10}(\mathrm{EDC}(t,\phi,\theta)). \tag{5}$$

In an isotropic sound field, the energy coming from any incident direction is equivalent to a mean calculated over all directions. As such, the next step is to calculate this mean for a chosen set of N directions (ϕ_i,θ_i) as

$$\overline{\mathrm{EDC}}_{\mathrm{dB}}(t) = \frac{1}{N}\sum_{i=0}^{N-1}\mathrm{EDC}_{\mathrm{dB}}(t,\phi_i,\theta_i). \tag{6}$$

We define the EDD for each direction as the deviation from the isotropic mean (Alary *et al.*, 2019a):

$$\mathrm{EDD}(t,\phi,\theta) = \mathrm{EDC}_{\mathrm{dB}}(t,\phi,\theta) - \overline{\mathrm{EDC}}_{\mathrm{dB}}(t). \tag{7}$$

The EDD values represent how much energy remains in the decay at a given time and direction relative to the $\overline{\mathrm{EDC}}_{\mathrm{dB}}$. Keeping in mind the smoothing caused by the beamformer, the range of the deviation itself is an important piece of information in the EDD.

The EDD can also be used to analyze the decay characteristics in the frequency domain simply by replacing the EDC with frequency-dependent EDR curves in Eq. (6) for a set of center-frequency bands ω. We obtain a frequency-dependent mean EDR from

$$\overline{\mathrm{EDR}}_{\mathrm{dB}}(t,\omega) = \frac{1}{N}\sum_{i=0}^{N-1}\mathrm{EDR}_{\mathrm{dB}}(t,\omega,\phi_i,\theta_i), \tag{8}$$

from which we can calculate the frequency-dependent EDD using

$$\mathrm{EDD}(t,\omega,\phi,\theta) = \mathrm{EDR}_{\mathrm{dB}}(t,\omega,\phi,\theta) - \overline{\mathrm{EDR}}_{\mathrm{dB}}(t,\omega). \tag{9}$$

Due to the inherent limitations of spherical microphone arrays, a bandpass filter should be applied to limit the frequency range of the analysis (Rafaely, 2005).

To illustrate how different directional decay times may affect the anisotropy, an artificial signal was generated using a set of decaying Gaussian white noise signals distributed to a set of points around a sphere. Individual signals were created using a set of direction-dependent decay times $T_{60}(\phi,\theta)$ distributed using a cardioid pattern (Fig. 2). Each signal was mixed with another noise sequence of fixed amplitude (–60 dB), representing the noise floor, and was encoded into fourth-order Ambisonics for each angle pair (ϕ_i,θ_i).

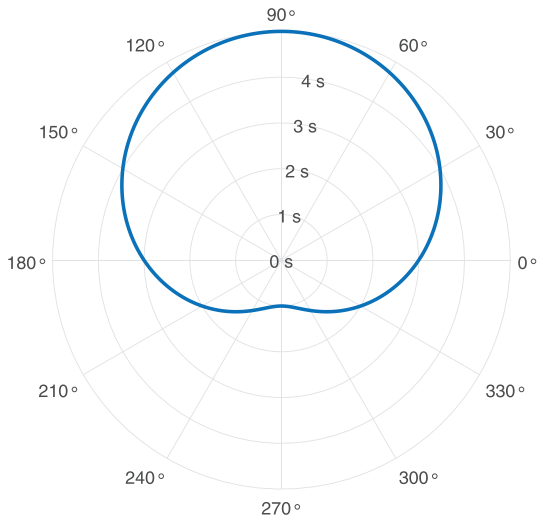


FIG. 2. (Color online) $T_{60}(\phi,0^\circ)$ of an artificial test signal, showing its distribution on the azimuthal plane. The signal was generated to approximate a cardioid pattern.

In Fig. 3, we can see the resulting EDD on the horizontal plane. Darker areas represent larger deviations to $\overline{\mathrm{EDC}}_{\mathrm{dB}}$, while white represents no deviation. In the online version, the red color highlights areas where more energy remains in the decay, therefore showing directions with a longer T_{60}, while the blue shows the opposite.

B. Interaural-energy-decay deviation (IEDD)

The SIR can also be encoded to a binaural signal using a set of head-related transfer functions (HRTFs) to yield a binaural room impulse response (BRIR), which can be useful for an objective measure closer to human perception. Through this, the direction-dependent characteristics are collapsed into frequency-dependent perceptual attributes for a fixed orientation of the sound field. Using the BRIR, we can compute the EDR of each binaural channel separately.

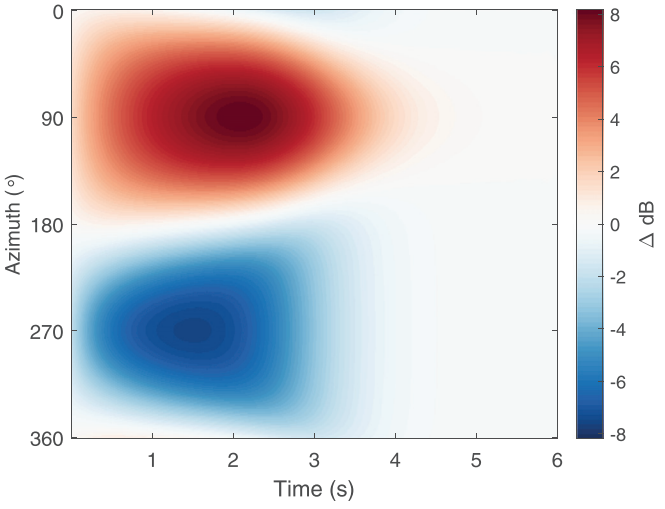


FIG. 3. (Color online) EDD of the artificial signal on the azimuthal plane. Darker areas represent larger deviations to $\overline{\mathrm{EDC}}_{\mathrm{dB}}$, while white represents no deviation.

https://doi.org/10.1121/10.0004770

Here, we no longer have a meaningful $\overline{\mathrm{EDC}}$ to use as reference. Instead, we look at the IEDD between the two binaural channels, which can be calculated as

$$\mathrm{IEDD}(t,\omega) = \mathrm{EDR}^{\mathrm{L}}_{\mathrm{dB}}(t,\omega) - \mathrm{EDR}^{\mathrm{R}}_{\mathrm{dB}}(t,\omega), \tag{10}$$

where the $\mathrm{EDR}^{\mathrm{L}}_{\mathrm{dB}}$ is computed from the left channel of the BRIR and the $\mathrm{EDR}^{\mathrm{R}}_{\mathrm{dB}}$ from the right channel. For an overview of the energy deviation per frequency, we calculate the root mean square over the time axis

$$\mathrm{IEDD}(\omega) = \sqrt{\frac{1}{T}\sum_{t=0}^{T-1}\mathrm{IEDD}(t,\omega)^2}. \tag{11}$$

C. Results

Several impulse responses were analyzed using the EDD method. This section details the objective results from four recorded SIRs. The SIRs were all recorded in fourth-order Ambisonics using the 32-capsule Eigenmike® microphone array (mhacoustics, 2020). The same SIRs are also used in the perceptual evaluation detailed in Sec. IV.

For visualization purposes, we only show the azimuthal plane in the following result plots, which corresponds to the 0° azimuthal plane in the recording setup. Here, $\overline{\mathrm{EDC}}_{\mathrm{dB}}$ was calculated from the signals taken from the azimuthal plane as well, meaning it is not an average over the full sphere. An average taken from points around the sphere would lead to poor visualization if a dominant direction were located outside the azimuthal plane. Video files containing the full spherical EDD analysis are included as supplementary material.[1]

In the following descriptions, the mid-frequency T_{60}^{mid} is defined as the mean reverberation time for all directions between the 500 Hz and 1 kHz octave bands.

1. Athénée Theatre

Figure 4(a) shows the EDD taken on the azimuthal plane from a SIR captured at the Athénée Theatre in Paris, France. This theatre is a 550-seat late 19th-century Italian-style hall. The loudspeaker and microphone were 12.4 m apart, and the mixing time was estimated to be 173 ms. This particular measurement was made with the source loudspeaker on the open stage and the receiving spherical microphone array on the far audience-left side of the hall next to a doorway opened onto an adjacent garden, about halfway down the orchestra level. As such, the measurement was deliberately set up to have strong anisotropic characteristics. In the EDD figures, we observe strong energy centered around 90°, and the T_{60}^{mid} is 1.50 s.

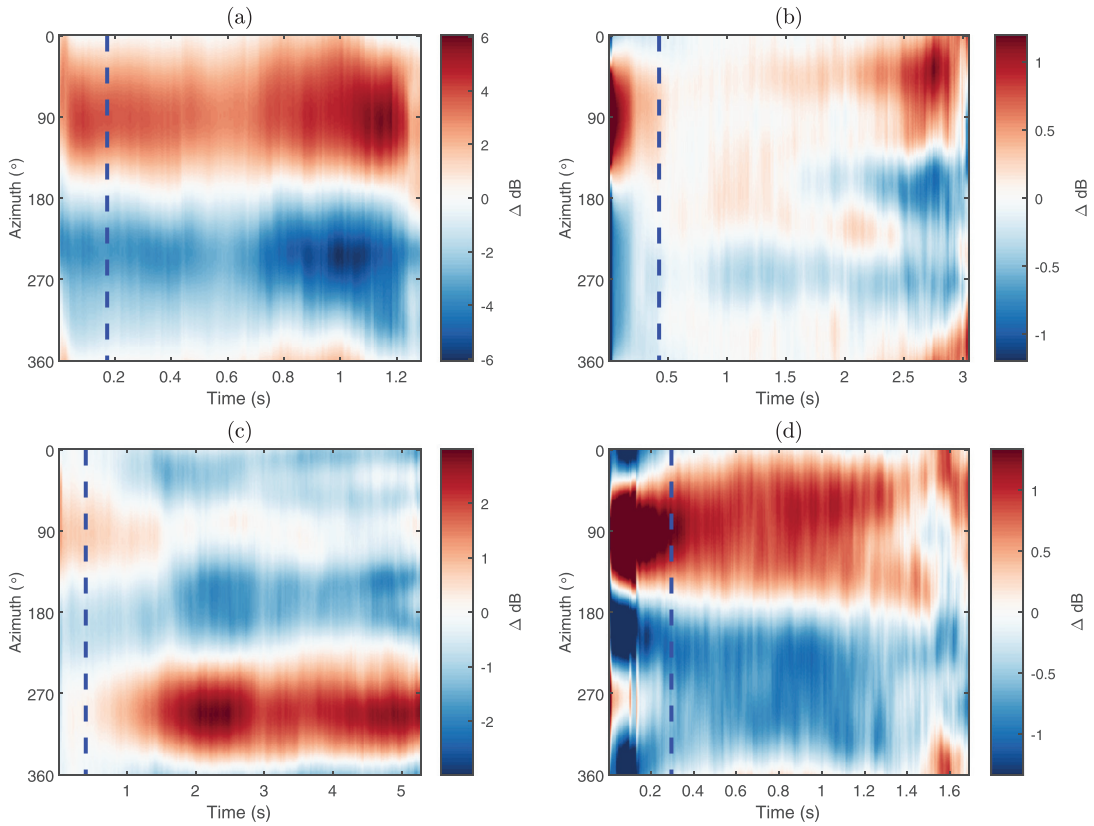


FIG. 4. (Color online) EDD analysis on the lateral plane of four halls: (a) the Athénée Theatre, (b) the Church of Saint Eustache, (c) the Staatliche Kunsthalle art museum, and (d) the Badisches Staatstheater. In the online version, the red areas represent the dominant directions of late reverberation. The dashed vertical blue line represents the estimated mixing time.

2. Church of Saint Eustache

In Fig. 4(b), the EDD represents the azimuthal plane from a SIR captured at the Church of Saint Eustache in Paris, France. The Church of Saint Eustache is a large 17th-century Gothic church. The church is approximately 100 m long, 40 m wide, and 30 m tall. In this SIR, the microphone was 33.9 m away from the loudspeaker, and the estimated mixing time was 428 ms. The measurement used here was captured with the source centered at the foot of the nave and the receiver centered on the steps leading from the crossing to the choir. A much smaller range of deviation is observed in the EDD here, with some frequency-dependent characteristics. The energy of the early reflections, before the mixing time, is concentrated around 90°, and the deviation range is very small during the first 2 s of late reverberation. The T_{60}^{mid} of the church is measured at 6.2 s.

3. Staatliche Kunsthalle art museum

In Fig. 4(c), the SIR was captured at the Staatliche Kunsthalle art museum in Karlsruhe, Germany in the museum's permanent exhibition space, with the source in one display room and the receiver through a large doorway and around the corner in another, resulting in an indirect coupled volume configuration. A deviation range of approximately 8 dB is present in the broadband analysis along with dominant energy throughout the decay centered around 290° for an estimated mixing time of 397 ms. In Fig. 5, we observe a second dominant direction emerging near 90° at 4000 Hz. Here, the sound source and the microphone were 16 m apart, but there was no direct path between the sound source and the receiver, since they were both on a different side of the coupled volume, and the microphone was located more than 2 m away from the closest wall. In this space, the T_{60}^{mid} is 4.2 s.

4. Badisches Staatstheater

The SIR in Fig. 4(d) was recorded at the Badisches Staatstheater in Karlsruhe, Germany. The Staatstheater is a modern 1000-seat opera and theatre hall (opened in 1975) with wood paneling on concrete walls and an asymmetric layout. The measurement used here was made with both source and receiver at the orchestra level and centered with respect to the stage. The stage area was closed off with an iron curtain, and the orchestra pit was covered with flooring, thereby removing any potential coupled spaces. The source was placed in one of the last rows, and the receiver was approximately 6.2 m away, centered toward the stage. In this SIR, the EDD analysis of the azimuthal plane yields clear dominant energy centered at 90°, which is stable across frequencies and time. In this hall, the deviation range is approximately 4 dB, and the T_{60}^{mid} is 1.7 s. Early in the decay, more energy is observed toward 275°, but it quickly dissipates after the mixing time, which was estimated to be 295 ms.

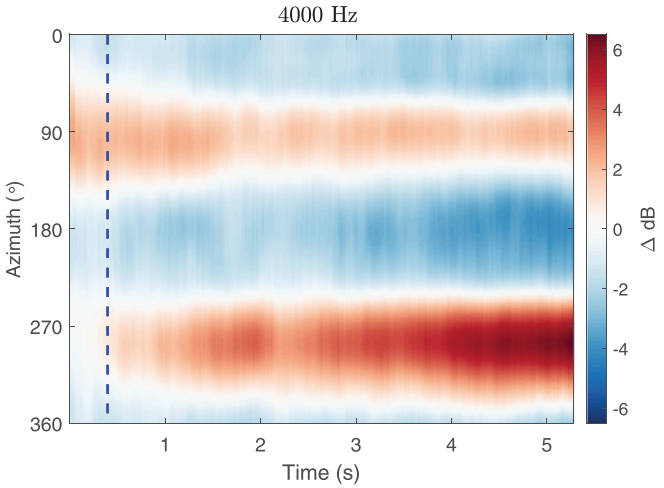


FIG. 5. (Color online) EDD analysis of the Staatliche Kunsthalle art museum at 4 kHz, cf. Fig. 4(c).

D. Interaural EDD

In Fig. 6, each of the above SIRs were also analyzed using the IEDD(ω) method described by Eq. (11). Each curve shows the spectral deviations that occur in the energy decay between the left and right binaural channels, which illustrates the perceptual attribute when a listener faces a specific direction. In Fig. 6(a), the listener faces (0°, 0°) and in Fig. 6(b) (135°, 0°). Each BRIR was encoded using the Ambisonics-to-binaural plugin included in the SPARTA suite (McCormack and Politis, 2019) with the default set of HRTFs. The differences between Figs. 6(a) and 6(b) illustrate the impact of head orientation on the IEDD characteristics due to the spectral envelopes of the direction-dependent HRTF filters. Note that the SIRs used here, both normal and rotated, are the same ones used in the perceptual study detailed in Sec. IV.

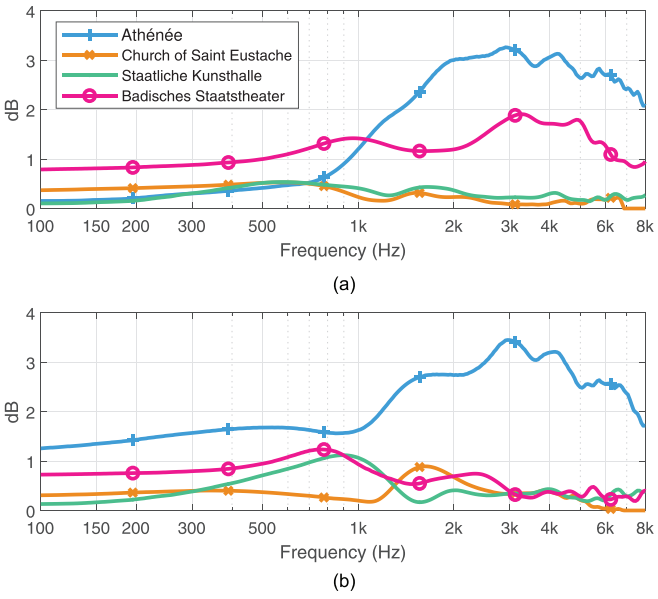


FIG. 6. (Color online) IEDD(ω) curves of the four analyzed SIRs in (a) the non-rotated and (b) the 135° rotated sound field. The dB values represent the spectral deviation, averaged over time, between the left and the right ear.

https://doi.org/10.1121/10.0004770

Through this analysis, the full perceived sound field is analyzed, and since we average over time, a smoothing of the values occurs. For these reasons, the analysis yields different information than the EDD analysis of the azimuthal plane. Nonetheless, the Athénée Theatre still has the strongest attributes, whereas the Church of Saint Eustache has the lowest. The analysis of the Staatliche Kunsthalle art museum has a slight peak centered around 800 Hz, whereas the Badisches Staatstheater has a higher frequency-dependent deviation above 3 kHz.

IV. SUBJECTIVE EVALUATION

Although the proposed objective evaluation method demonstrates anisotropy in the four cases detailed above, verifying that this anisotropy is in fact audible is important. Understanding the audibility of an anisotropic sound field is also crucial to help determine when this anisotropy is important in reproduction. Since the key assumption in multichannel reverberation is that the decaying sound field is perceived to be isotropic after the mixing time, the perceptual test is constrained to the audibility of the sound field after this mixing time. The following perceptual study assesses the capacity of a listener to detect the perceptual cues that arise from a rotation on the azimuthal plane of a SIR. To abstract the perception of the early specular reflections, the rotation is only performed after the mixing time, which means that the early reflections are static throughout the experiment.

A. Overview

To create a new set of SIRs containing a rotated sound field, a rotation of 135° on the azimuthal plane was performed in the SHD using a Euler rotation matrix. The beginning of the unrotated version of each of the chosen SIRs was then mixed together with the corresponding rotated late parts using a short cross-fade of 10 samples on every Ambisonics channel to transition between the early part and late reverberation part at t_{mix}. No test subject reported any audible artefacts from the cross-fade during the perceptual study.

B. Hypotheses

The subjective evaluation was designed to verify the following hypotheses:

$\mathcal{H}_0$: the subjects cannot identify when the sound field is rotated in an artificial isotropic SIR,
$\mathcal{H}_1$: the subjects can differentiate between the rotated and non-rotated recorded SIRs,
$\mathcal{H}_2$: the identification rate is positively linked with the maximum range of values in the IEDD(ω),
$\mathcal{H}_3$: stimuli with a broader frequency spectrum are easier to identify.

C. Apparatus and calibration

The perceptual study was conducted in the facilities of the Acoustics Lab of Aalto University, located in Espoo, Finland (Fig. 7). The anechoic chamber is an extremely silent space, as its A-weighted background noise level is −2.1 dB when the loudspeakers are turned off and 11.6 dB when the loudspeakers are turned on (Kuusinen and Lokki, 2020). The room has 350-mm thick absorbent material on every surface and meets the free-field conditions from 50 Hz upward, which satisfies the requirements of ISO 3745:2003 (2003). The inside of the room is approximately 5 m wide, 5 m long, and 5 m high with a metal grid floor suspended 1 m above the bottom.

The loudspeaker array consists of 37 Genelec Ones 8331 A speakers, which are uniformly distributed on five circular rings with one extra speaker overhead, as illustrated in Fig. 8. The computer was connected to an RME ADI-6432 audio interface via a RME MADIface XT external soundcard module. The RME ADI-6432 sends the 37-channel audio signals to the loudspeakers in the anechoic multichannel room.

The loudspeaker array was calibrated using the calibration software recommended by Genelec (Iisalmi, Finland), GLM 3. The calibration procedure for the perceptual study consisted of measuring sine sweeps from all individual loudspeakers at the listening position. The system then optimized the loudspeaker levels and delays to ensure balanced levels and synchronized times of arrival from every loudspeaker to the listening position. Also, as part of the calibration procedure, the frequency responses of the loudspeakers were analyzed, and the main peaks in their magnitude response were equalized to ensure a neutral sound reproduction.

D. Spatial coherence

The spatial coherence introduced by the Ambisonics decoder was evaluated to ensure that any artefacts introduced in the decoding phase did not interfere with the perceptual study. For this purpose, a set of fully incoherent and isotropic signals was produced using white Gaussian noise

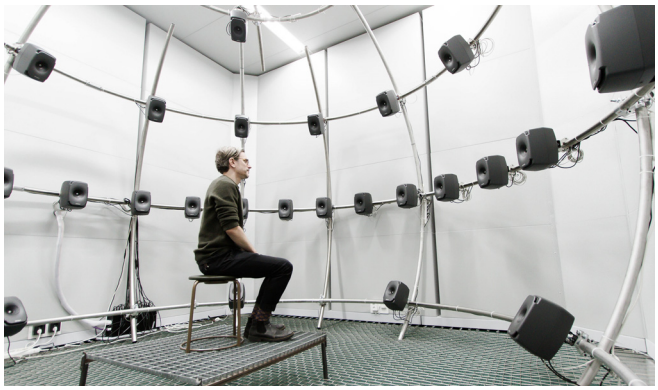


FIG. 7. (Color online) Picture of the multichannel reproduction room used in the perceptual study.

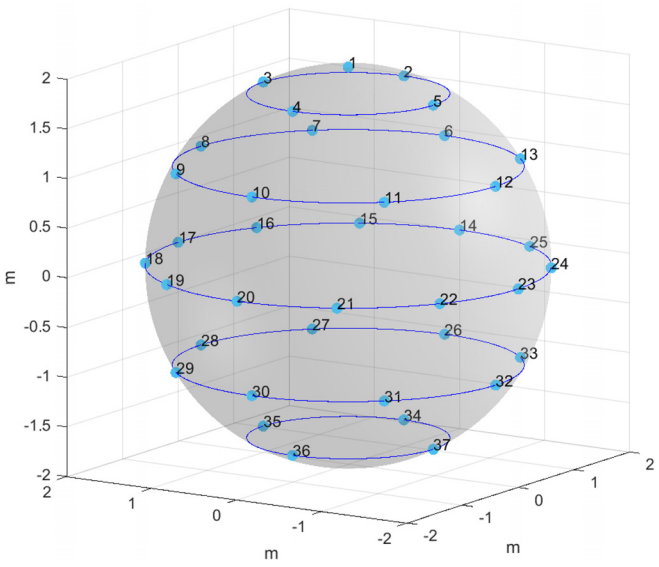


FIG. 8. (Color online) Channel distribution of the loudspeaker array used for reproduction in the perceptual study, cf. Fig. 7. Channel one is directly above the listener.

with a common decay envelope. Individual signals were distributed to a specific point on a t-design spherical grid of 840 points before encoding them to fourth-order Ambisonics. The encoded signal was in turn decoded to the loudspeaker array configuration used in the perceptual study.

Figure 9 shows the coherence matrix between the different output channels. Here, 1 represents two fully coherent channels and 0 two fully incoherent ones. The diagonal line represents the coherence between each channel and itself, which is always 1. Figure 9 indicates that a small amount of coherence is introduced by this decoding, which is expected in Ambisonics signals.

To obtain the results in Fig. 10, the coherence matrix measured before and after applying a rotation of 135° is

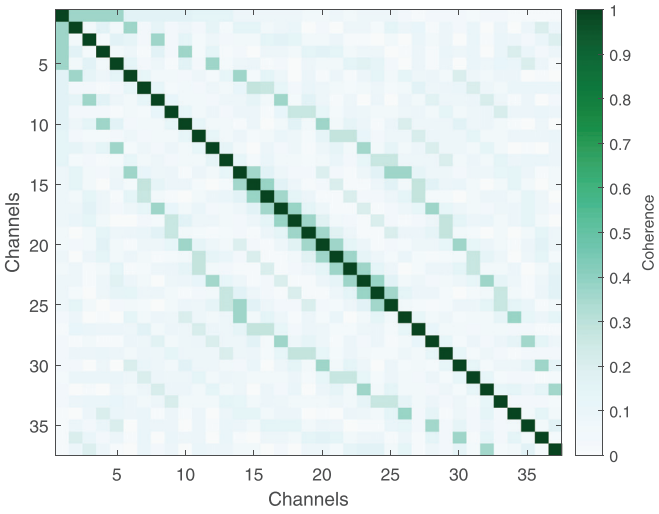


FIG. 9. (Color online) Coherence between channels after encoding an artificial signal to fourth-order Ambisonics and decoding it to the specified loudspeaker array.

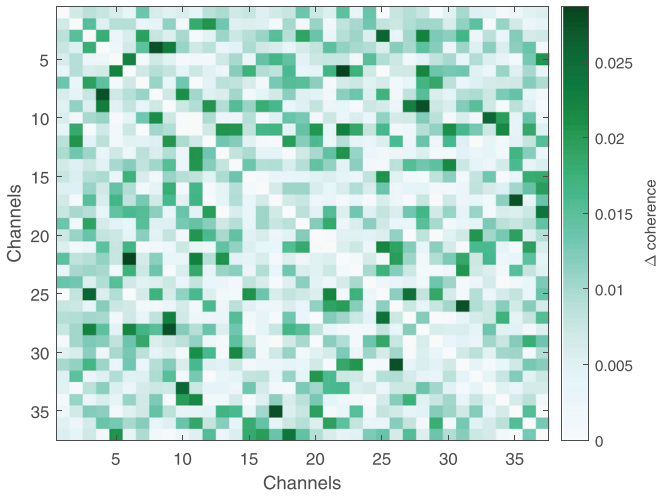


FIG. 10. (Color online) Differences in coherence between a 135° rotation of the artificial signal on the azimuthal plane and the original, non-rotated signal. White (0) means that the coherence is exactly the same. Here, the range of values is very small and, as such, perceptually negligible.

subtracted from the isotropic test signal. The maximum difference in coherence is less than 0.025, which is negligible. This suggests that the loudspeaker array is well distributed and that any coherence introduced in the decoding stage will be consistent throughout the subjective evaluation.

Generally, we expect an impulse response to have very low coherence between channels after t_{mix} due to the large amount of uncorrelated plane waves coming from all directions. To validate this hypothesis, we measured the difference in coherence between the original and rotated impulse response used in the listening test. These results are not included here, since they are all <0.1 and hence negligible and similar to the synthetic example shown in Fig. 10.

E. Stimuli

The stimuli were chosen to be varied and to represent usual broadcasting sounds as specified by the International Telecommunication Union (2015). The chosen samples consist of recordings of a trumpet, a male voice, percussion, and a guitar. All stimuli were recorded in acoustically dry conditions. A synthetic signal was also used, serving as a reference stimulus. This synthetic stimulus is based on a pink impulse, a linear-phase signal with a spectrum corresponding to $H(\omega) = 1/\sqrt{\omega}$ (Liski *et al.*, 2018). The goal was to assess whether its broader frequency range yields a higher identification rate than the natural sounds by exciting more room modes in a SIR ($\mathcal{H}_3$).

All the above stimuli were convolved with the four SIRs evaluated in Sec. III C as well as the synthetic isotropic signal used in Sec. IV D, which served as the anchor to confirm the null hypothesis $\mathcal{H}_0$. The same samples were also convolved with the rotated version of the same SIRs following the procedure detailed in Sec. IV A. The resulting Ambisonics files were then decoded for the loudspeaker array layout.

https://doi.org/10.1121/10.0004770

F. Participants

For the perceptual study, 18 audio researchers participated, all with prior experience in spatial audio perceptual evaluation and no reported hearing impairments. The mean age of the subjects was 33. The results from two of the first participants were discarded, bringing the total down to 16. These participants received different instructions from the others, and they had poor results with the reference anisotropic SIR, which was correctly identified by all the other subjects.

G. Method

The perceptual study follows the guidelines for assessing small impairments in audio systems proposed in ITU-R BS.1116–3 (International Telecommunication Union, 2015). With this method, we assessed the ability of a listener to detect small differences occurring when a decaying sound field is rotated. A reference signal was presented to the listeners along with two blind stimuli, and their task was to identify the reference from the two stimuli.

The experiment was implemented using MAX (Cycling'74, San Francisco, CA). The main interface, shown in Fig. 11, was displayed on a tablet with which the test subject could select to play and stop either the reference or the two blind stimuli. The subjects could switch between stimuli at any time during playback, but no reverberation tail was heard if a stimulus was stopped before the end, since the stimuli were encoded in advance. An equal-power crossfade of 50 ms was applied to transitions between the stimuli. The subjects were instructed to adjust the volume to a comfortable level. All the SIRs were captured beyond the critical distance, meaning that the reverberant part of the signal contained more energy than the direct sound.

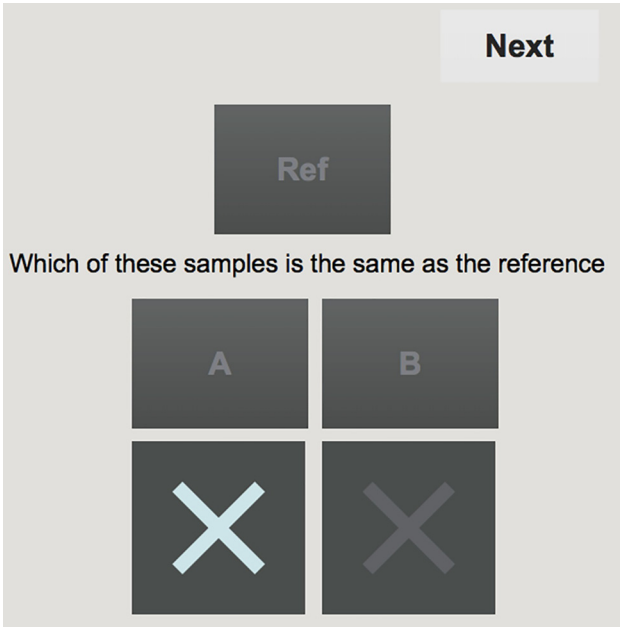


FIG. 11. (Color online) Interface used in the perceptual study. Pressing the "Ref," "A," or "B" buttons switches playback or stops it if it was already playing.

The subjects were instructed to pay special attention to the directions perceived as dominant during the late reverberation. The direction of the direct sound varied from one SIR to another but was consistent between stimuli of the same SIR, since the rotation was applied after the mixing time. The subjects were encouraged to rotate their body while remaining seated in the sweet spot of the room to vary the listening angle and minimize the impact of direction-dependent binaural cues, such as the cone of confusion, as illustrated by Fig. 6. Each subject was presented with the same 50 stimuli in random order.

H. Results of perceptual study

Figure 12 shows the results for individual test subjects, identified with light circles. These individual results are all multiples of 10%, since each SIR was presented ten times to each subject (five stimuli with one repetition). The horizontal dashed line represents the confidence line for a set of M trials, which is calculated using

$$M/2 + \sqrt{M}. \tag{12}$$

The coherence between the various output channels, before and after the rotation, shows very low correlation between channels, as mentioned in Sec. IV D. Since $\mathcal{H}_0$ states that a rotation in the late part of an isotropic sound field with low coherence is not identifiable, this suggests that the inter-channel cross correlation was not a key factor in discriminating between the stimuli. The results from Athénée, Kunsthalle, and Staatstheater all support $\mathcal{H}_1$, but the rotation of the sound field in the Church of Saint Eustache, which also has the smallest range of IEDD(ω), was not identified in a statistically significant way, as its

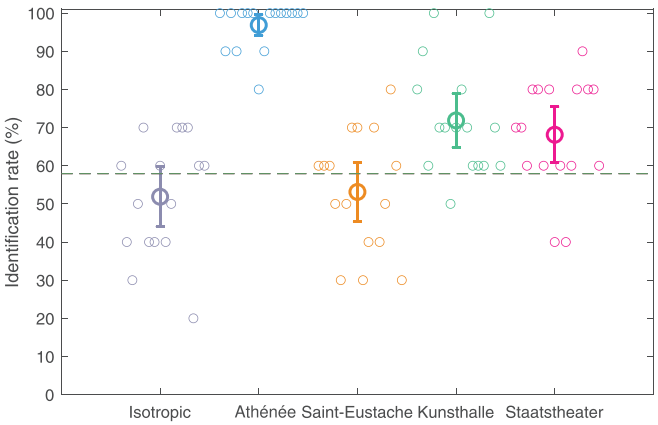


FIG. 12. (Color online) Listening-test results for each participant are represented by thin circles that are spread over the horizontal axis for visibility. The average identification rate of each SIR is presented as a thick circle, with the line segment representing its 95% confidence interval. The confidence line for $M = 160$ trials is indicated with a horizontal dashed line. The results show that the rotated sound field was identified in a statistically significant way for the SIRs recorded at Athénée, Kunsthalle, and Staatstheater.

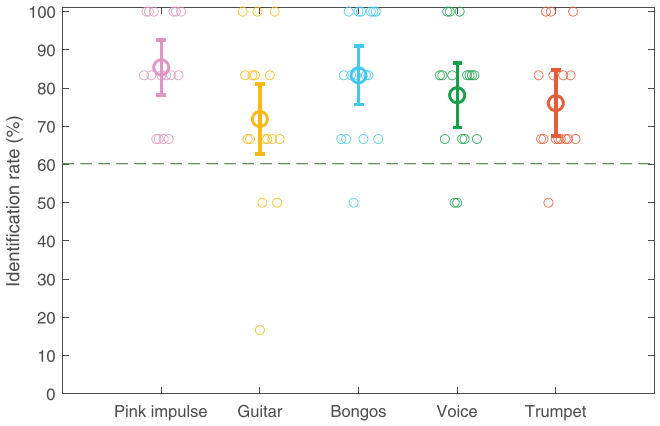


FIG. 13. (Color online) Listening-test results per stimulus, using only the SIRs that were well identified in Fig. 12 (Athénée, Kunsthalle, and Staatstheater). The results per participant are represented by thin circles. The average of each stimulus is shown with a thick circle with a 95% confidence interval. The horizontal dashed line is the confidence line for $M = 96$ trials. These results demonstrate that no statistically significant differences were noted between the stimuli, since their confidence intervals overlap.

average result is lower than the confidence line in Fig. 12. Athénée, which has the highest identification rate in Fig. 12, has also the largest values of IEDD(ω) (see Fig. 6), which is consistent with $\mathcal{H}_2$.

The result labeled "Isotropic" in Fig. 12 refers to the artificial signal created to be fully isotropic, as detailed in Sec. IV D. These results indicate that beyond a threshold close to IEDD$(\omega) > 1$ dB, the reproduction of direction-dependent decays is necessary for an accurate reproduction of the sound field. In the case of the Staatliche Kunsthalle art museum, an IEDD(ω) of 1.1 dB around 850 Hz was sufficient to differentiate the stimuli.

Figure 13 shows the results separately for each stimulus, using the results from the three identifiable rotated SIRs (Athénée, Kunsthalle, and Staatstheater). The results indicate a slight increase in detection rate using the pink impulse, but overall, no statistically significant differences occur in perception between each stimulus, which contradicts $\mathcal{H}_3$. In other words, the results demonstrate that the anisotropic late reverberation is audible with all tested sounds.

I. Discussion

We can see that although the reverberated signal is well decorrelated in every direction, some directions may still exhibit some energy deviation beyond the mixing time. Since this anisotropy in late reverberation is also perceivable in the selected examples, these results suggest the importance of considering direction-dependent characteristics in the decay. The positive link between the objective measures and the perceptual detection rate in our results suggests that this analysis framework is suitable to assess the perception of anisotropy in specific cases. While more work remains to establish the specific spectro-temporal threshold of these characteristics, the framework detailed in this article should help future studies on this topic.

V. CONCLUSION

In conclusion, we introduced a framework to analyze and assess the anisotropic features of SIRs, both objectively and subjectively. The proposed objective measure can highlight direction-dependent characteristics in a decaying sound field and can serve as an analysis tool to estimate the energy deviation in a perceived binaural sound field. A subjective evaluation method was proposed to assess our capacity to hear these anisotropic characteristics by rotating the sound field after the mixing time.

The perceptual study performed with this method demonstrated a correspondence between the direction-dependent deviation in the energy decay and the detection rate between stimuli. Although more work is necessary to identify a precise perceptual threshold for these characteristics, our experiment found that an IEDD of 1.1 dB around 850 Hz was sufficient to identify the rotated stimulus. Therefore, the results detailed in this paper suggest that reproducing the direction-dependent characteristics in late reverberation is more important than previously thought and that special attention should be paid to the amount of direction-dependent deviations in the energy decay for the accurate reproduction of spatial sound.

Future work includes studying specific factors influencing the perception of anisotropic characteristics, such as the visual appearance of the space, overall volume, and the proximity between the source and the listener, that could produce a masking effect in this context.

ACKNOWLEDGMENTS

Part of this work was conducted during B.A.'s research visits to the Institut de Recherche et Coordination Acoustique/Musique (IRCAM) (UMR STMS IRCAM-CNRS-Sorbonne Université), Paris in October–December 2018 and September 2019, funded by the Foundation for Aalto University Science and Technology. This work was funded in part by the Academy of Finland (ICHO project, Aalto University Project No. 13296390) and by the RASPUTIN project (Grant No. ANR-18-CE38-0004), and it is part of the activities of the Nordic Sound and Music Computing Network—NordicSMC (NordForsk Project No. 86892). Additional support for P.M. was provided through the doctoral research grant from the École doctorale Informatique, Télécommunications, et Électronique (EDITE) at Sorbonne Université. The authors would like to thank Archontis Politis and Olivier Warusfel for fruitful discussions on the perceptual study as well as Augustin Muller and Pedro Garcia-Velazquez for their extensive SIR measurements made during their Artistic Research Residency at IRCAM. M.N. and V.V. contributed equally to the supervision of this work.

[1]See supplementary material at https://www.scitation.org/doi/suppl/10.1121/10.0004770 for a full EDD analysis of the spherical sound field. The file list includes the video rendering of the EDD of the Athénée Theatre (SuppPubmm1.avi), the Church of Saint Eustache (SuppPubmm2.avi), the Staatliche Kunsthalle (SuppPubmm3.avi), and the Badisches Staatstheater (SuppPubmm4.avi).

Ahonen, J., and Pulkki, V. (**2009**). "Diffuseness estimation using temporal variation of intensity vectors," in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, October 18–21, New Paltz, NY, pp. 285–288.

Alary, B., Massé, P., Välimäki, V., and Noisternig, M. (**2019a**). "Assessing the anisotropic features of spatial impulse responses," in *Proceedings of the EAA Spatial Audio Signal Processing Symposium*, September 6–7, Paris, pp. 43–48.

Alary, B., and Politis, A. (**2020**). "Frequency-dependent directional feedback delay network," in *Proceedings of the IEEE ICASSP-2020*, May 4–8, Barcelona, Spain, pp. 176–180.

Alary, B., Politis, A., Schlecht, S. J., and Välimäki, V. (**2019b**). "Directional feedback delay network," J. Audio Eng. Soc. **67**(10), 752–762.

Balachandran, C. G., and Robinson, D. W. (**1967**). "Diffusion of the decaying sound field," Acta Acust. **19**(5), 245–257.

Berzborn, M., and Vorländer, M. (**2018**). "Investigations on the directional energy decay curves in reverberation rooms," in *Proceedings of Euronoise 2018*, May 27–31, Heraklion, Greece, pp. 2005–2010.

Blesser, B. A. (**2001**). "An interdisciplinary synthesis of reverberation viewpoints," J. Audio Eng. Soc. **49**(10), 867–903.

D'Antonio, P., Jeong, C. H., and Nolan, M. (**2018**). "Design of a new test chamber to measure the absorption, diffusion, and scattering coefficients," J. Acoust. Soc. Am. **144**(3), 1814–1814.

Epain, N., and Jin, C. T. (**2016**). "Spherical harmonic signal covariance and sound field diffuseness," IEEE/ACM Trans. Audio Speech Lang. Proc. **24**(10), 1796–1807.

Fliege, J., and Maier, U. (**1996**). "A two-stage approach for computing cubature formulae for the sphere," in *Mathematik 139T* (Universitat Dortmund, Dortmund, Germany).

Gerzon, M. A. (**1972**). "Synthetic stereo reverberation: Part II," Studio Sound **14**, 24–28.

Gerzon, M. A. (**1975**). "Recording concert hall acoustics for posterity," J. Audio Eng. Soc. **23**(7), 569–571.

Götz, P., Kowalczyk, K., Silzle, A., and Habets, E. A. P. (**2015**). "Mixing time prediction using spherical microphone arrays," J. Acoust. Soc. Am. **137**(2), EL206–EL212.

Gover, B. N., Ryan, J. G., and Stinson, M. R. (**2002**). "Microphone array measurement system for analysis of directional and spatial variations of sound fields," J. Acoust. Soc. Am. **112**(5), 1980–1991.

Gover, B. N., Ryan, J. G., and Stinson, M. R. (**2004**). "Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array," J. Acoust. Soc. Am. **116**(4), 2138–2148.

Guski, M., and Vorländer, M. (**2014**). "Comparison of noise compensation methods for room acoustic impulse response evaluations," Acta Acust. united Acust. **100**(2), 320–327.

International Telecommunication Union (**2015**). "Recommendation ITU-R BS. 1116-3: Methods for the subjective assessment of small impairments in audio systems," International Telecommunication Union, Geneva, Switzerland.

ISO 3745:2003 (**2003**). "Acoustics—Determination of sound power levels of noise sources using sound pressure—Precision methods for anechoic and hemi-anechoic rooms" (International Organization for Standardization, Geneva, Switzerland).

Jarrett, D. P., Thiergart, O., Habets, E. A. P., and Naylor, P. A. (**2012**). "Coherence-based diffuseness estimation in the spherical harmonic domain," in *Proceedings of the IEEE 27th Convention of Electrical and Electronics Engineers in Israel*, November 14–17, Eilat, Israel, pp. 1–5.

Jot, J.-M. (**1992**). "An analysis/synthesis approach to real-time artificial reverberation," in *Proceedings of the IEEE ICASSP-92*, March 23–26, San Francisco, CA, pp. 221–224.

Jot, J.-M., Cerveau, L., and Warusfel, O. (**1997**). "Analysis and synthesis of room reverberation based on a statistical time-frequency model," in *Proceedings of the Audio Engineering Society 103rd Convention*, September 26–29, New York.

Jot, J.-M., and Chaigne, A. (**1991**). "Digital delay networks for designing artificial reverberators," in *Proceedings of the Audio Engineering Society 90th Convention*, February 19–22, Paris.

Karjalainen, M., Antsalo, P., Mäkivirta, A., Peltonen, T., and Välimäki, V. (**2002**). "Estimation of modal decay parameters from noisy response measurements," J. Audio Eng. Soc **50**(11), 867–878.

Kuttruff, H. (**2009**). *Room Acoustics*, 5th ed. (Taylor & Francis, London).

Kuusinen, A., and Lokki, T. (**2020**). "Recognizing individual concert halls is difficult when listening to the acoustics with different musical passages," J. Acoust. Soc. Am. **148**(3), 1380–1390.

Lachenmayr, W., Haapaniemi, A., and Lokki, T. (**2016**). "Direction of late reverberation and envelopment in two reproduced Berlin concert halls," in *Proceedings of the Audio Engineering Society 140th Convention*, June 4–7, Paris.

Lindau, A., Kosanke, L., and Weinzierl, S. (**2012**). "Perceptual evaluation of model- and signal-based predictors of the mixing time in binaural room impulse responses," J. Audio Eng. Soc. **60**(11), 887–898.

Liski, J., Mäkivirta, A., and Välimäki, V. (**2018**). "Audibility of loudspeaker group-delay characteristics," in *Proceedings of the Audio Engineering Society 144th Convention*, May 23–26, Milan, Italy.

Luizard, P., Katz, B. F. G., and Guastavino, C. (**2015**). "Perceptual thresholds for realistic double-slope decay reverberation in large coupled spaces," J. Acoust. Soc. Am. **137**(1), 75–84.

Massé, P., Carpentier, T., Warusfel, O., and Noisternig, M. (**2020a**). "Denoising directional room impulse responses with spatially anisotropic late reverberation tails," Appl. Sci. **10**(3), 1033.

Massé, P., Carpentier, T., Warusfel, O., and Noisternig, M. (**2020b**). "A robust denoising process for spatial room impulse responses with diffuse reverberation tails," J. Acoust. Soc. Am. **147**(4), 2250–2260.

McCormack, L., and Politis, A. (**2019**). "SPARTA & COMPASS: Real-time implementations of linear and parametric spatial audio reproduction and processing methods," in *Proceedings of the Audio Engineering Society International Conference on Immersive and Interactive Audio*, March 27–29, York, UK.

mhacoustics (**2020**). https://mhacoustics.com/ (Last viewed May 3, 2021).

Nolan, M., Berzborn, M., and Fernandez-Grande, E. (**2020**). "Isotropy in decaying reverberant sound fields," J. Acoust. Soc. Am. **148**(2), 1077–1088.

Nolan, M., Fernandez-Grande, E., Brunskog, J., and Jeong, C.-H. (**2018**). "A wavenumber approach to quantifying the isotropy of the sound field in reverberant spaces," J. Acoust. Soc. Am. **143**(4), 2514–2526.

Pierce, A. D. (**1974**). "Concept of a directional spectral energy density in room acoustics," J. Acoust. Soc. Am. **56**(4), 1304–1305.

Polack, J.-D. (**1993**). "Playing billiards in the concert hall: The mathematical foundations of geometrical room acoustics," Appl. Acoust. **38**(2), 235–244.

Prasad, D. K., Leung, M. K., Quek, C., and Cho, S.-Y. (**2012**). "A novel framework for making dominant point detection methods non-parametric," Image Vision Comput. **30**(11), 843–859.

Pulkki, V. (**2007**). "Spatial sound reproduction with directional audio coding," J. Audio Eng. Soc. **55**(6), 503–516.

Rafaely, B. (**2005**). "Analysis and design of spherical microphone arrays," IEEE Trans. Speech Audio Process. **13**(1), 135–143.

Rafaely, B. (**2015**). *Fundamentals of Spherical Array Processing* (Springer, New York).

Romblom, D., Guastavino, C., and Depalle, P. (**2016**). "Perceptual thresholds for non-ideal diffuse field reverberation," J. Acoust. Soc. Am. **140**(5), 3908–3916.

Sakuma, T., and Eda, K. (**2013**). "Energy decay analysis of non-diffuse sound fields in rectangular rooms," Proc. Mtgs. Acoust. **19**(1), 015138.

Schlecht, S. J., and Habets, E. A. P. (**2017**). "Feedback delay networks: Echo density and mixing time," IEEE/ACM Trans. Audio Speech Lang. Process. **25**(2), 374–383.

Schroeder, M. R. (**1962**). "Natural sounding artificial reverberation," J. Audio Eng. Soc. **10**(3), 219–223.

Schroeder, M. R. (**1965**). "New method of measuring reverberation time," J. Acoust. Soc. Am. **37**(3), 409–412.

Schroeder, M. R., and Logan, B. F. (**1961**). "'Colorless' artificial reverberation," J. Audio Eng. Soc. **9**(3), 192–197.

Välimäki, V., Parker, J. D., Savioja, L., Smith, J. O., and Abel, J. S. (**2012**). "Fifty years of artificial reverberation," IEEE Trans. Audio Speech Lang. Process. **20**(5), 1421–1448.

Xiang, N., Trivedi, U., and Xie, B. (**2019**). "Artificial enveloping reverberation for binaural auralization using reciprocal maximum-length sequences," J. Acoust. Soc. Am. **145**(4), 2691–2702.