OPEN ACCESS Freely available online
J. Ahrens and L. Müller, "Perceptually Transparent Binaural Auralization of Simulated Sound Fields," *J. Audio Eng. Soc.*, vol. 74, no. 3, pp. 118–129 (2026 Mar.). https://doi.org/10.17743/jaes.2022.0252.

Perceptually Transparent Binaural Auralization of Simulated Sound Fields

JENS AHRENS,* *AES Member*, **AND LEON MÜLLER,** *AES Student Member*
(jens.ahrens@chalmers.se) (leon.muller@chalmers.se)

Division of Applied Acoustics, Chalmers University of Technology, Gothenburg, Sweden

Contrary to geometric acoustics-based simulations where the spatial information is available in a tangible form, it is not straightforward to auralize wave-based simulations. A variety of methods have been proposed that compute the ear signals of a virtual listener with known head-related transfer functions from sampling either the sound pressure or the particle velocity (or both) of the simulated sound field. This article summarizes the most common binaural auralization methods with and without intermediate ambisonic representation of volumetrically sampled sound pressure or sound pressure and particle velocity sampled on spherical or cubical surfaces and presents a perceptual validation thereof. A triangular test ($N = 19$) confirmed that all evaluated grids resulted in a perceptually transparent auralization for the three tested sound incidence angles under reverberant conditions. Under anechoic conditions, only the high-density spherical and cubical surface grids lead to transparent auralization. All tested methods are available open source in the Chalmers Auralization Toolbox that accompanies this article.

Keywords: auralization, radiation, binaural reproduction

1 INTRODUCTION

Wave-based methods like finite element method, boundary element method, and finite-difference time-domain are becoming increasingly popular for room acoustic simulations, and simulations over the entire audible frequency range are possible. Contrary to geometric acoustics-based simulations where the spatial information is available in a tangible form in terms of, for example, incidence directions and spectral fingerprints of reflections [1–4], it is not straightforward to auralize wave-based simulations, and a variety of approaches have been proposed.

The output of wave-based simulations may be termed a *virtual sound field*, and the format that the output of any wave-based simulation method can be converted to is the sound pressure and/or particle velocity at a set of observation points, which is termed here *sampling points* or *sampling nodes*. Those auralization approaches that require sampling of a virtual sound field on an infinitesimal portion of space around the notional receiver point to be auralized may be termed *local*. Examples are [5], which uses virtual microphones for obtaining a spherical harmonic (SH) representation of the sampled sound field, and [6] and [7], which obtain SH representations via higher-order spatial derivatives. Local approaches require higher-order time integration, which exhibits the inconvenience that it produces a DC drift that is difficult to tame [8].

Nonlocal approaches include [9], which obtains SH coefficients by numerically evaluating a boundary integral of pressure and velocity, and [10–13], which sample a portion of space volumetrically. A comprehensive evaluation of [9] is not available, so present considerations are limited to methods like [10–12] that are based on sampling the virtual sound field in a portion of space of moderate size.

Some of the methods mentioned above work solely based on the sampled sound pressure of the simulated field, which can lead to ill-conditioning of the problem. The methods that are considered here are closely related to binaural rendering of microphone array recordings in the literature where this ill-conditioning is well known [14, 15]. The set of sampling nodes in the present case can be interpreted as a virtual microphone array. The main difference is that, contrary to real microphones, the sampling nodes are ideal and their number and placement are virtually unrestricted. Combining pressure and pressure gradient can greatly improve the conditioning of the problem [16, 14]. One incarnation of this combination are (virtual or actual) microphones with cardioid directivity [15, 17], which are a promising solution to the present problem [18].

The authors want to emphasize that this work does not evaluate acoustical simulations themselves. The purpose is to investigate how nongeometrical acoustics simulations can be auralized binaurally (i.e., how to most favorably

*To whom correspondence should be addressed, email: jens.ahrens@chalmers.se.

compute signals from the output of the simulation that would arise at a listener's ears if the listener were exposed to the simulated sound field). As it is likely that it will not be possible to achieve that the output of the auralization is numerically identical to the ground truth, the authors seek for *perceptually transparent* auralization (i.e., auralization that is perceptually indistinguishable from the ground truth in a direct comparison).

This investigation is not based on the output of an actual simulation framework. Rather, a representation of room acoustic responses is used, from which both a binaural ground truth signal and sampled sound field data can be computed, the auralization of which is evaluated against the binaural ground truth. Previous perceptual evaluations of auralization of nongeometric acoustics simulations were of limited scope in that only one auralization method/grid was investigated at a time, and they considered a limited frequency bandwidth of at maximum 5 kHz [13, 19, 20]. The potential for general perceptual transparency is therefore unclear.

The auralization solution that this article seeks is intended to be as general as possible, which means that it only considers solutions that (1) do not make assumptions on the sound field that is to be auralized and (2) can work on running signals. The motivation for requirement 1 is that the auralization is desired to be applicable to any simulated sound field. There are methods that were originally proposed in the field of auralization of microphone array measurements of room impulse responses, such as [21, 22] that parameterize the sound field, for example, into direct and diffuse components. This can indeed provide a benefit over nonparametric methods, especially when the number of sampling nodes is small. Yet, there is uncertainty in the output from parametric methods if nonstandard sound fields are considered such as sound fields outdoors or in small compartments like car cabins, where it is not straightforward to differentiate direct sound from early reflections, or if occlusion of the direct sound occurs.

Linear and time-invariant (LTI) sound field simulations can be expressed in terms of impulse responses. While this is likely to be the most common application area of auralization today, requirement 2 is still formulated so that also dynamic scenarios and scenarios that comprise nonlinearities such simulations of electroacoustic systems can be handled. Parametric rendering methods have also been proposed for running signals [23, 24]. Those are excluded from the present investigation for the reasons stated above.

For convenience, this investigation is performed based on simulated room impulse responses. This does not limit the generalizability of the results as all methods can be straightforwardly applied to running signals without modification.

2 AURALIZATION FRAMEWORK

What is common to all auralization methods that are considered here is that the auralization itself is modelled as an LTI system and is represented by a set of impulse responses. The set of impulse responses essentially represents the multiple-input and multiple-output (MIMO) transfer path from the sampling nodes to the ears of the listener. Refer to Fig. 1 for an illustration. The transfer path therefore comprises the head-related transfer function (HRTF) of the listener as well as their head orientation. Once the impulse responses have been computed, carrying out the auralization consists in performing the required convolutions. This process if often termed *rendering* in spatial audio. The result of the rendering is the ear signals that would arise at the listener's ears if they were exposed to the virtual sound field with a given head orientation.

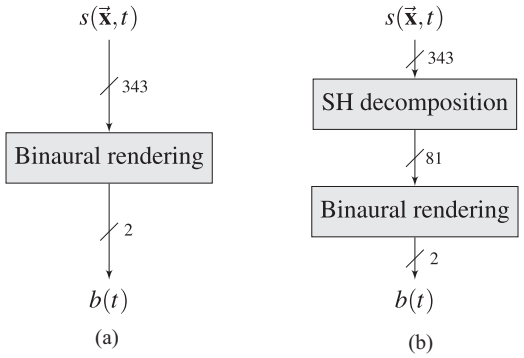


Fig. 1. Block diagram of the signal processing pipeline for direct auralization (a) and ambisonic auralization (b). $s(\vec{\mathbf{x}}, t)$ are signals at the sampling nodes $\vec{\mathbf{x}}$, and $b(t)$ is the binaural output. Each of the gray blocks represents a MIMO convolution. In these examples, 343 sampling nodes are assumed as well as an ambisonic representation of order $N = 8$, which produces an ambisonic signal with 81 channels. In either case, the binaural rendering stage takes the instantaneous head orientation of the listener into account.

The available methods may be divided in two categories: Methods that are termed here *direct methods* obtain the binaural output signals directly from the signals at the sampling nodes. Methods that are termed here *ambisonic methods* produce a representation of the sampled sound field in terms of SH basis functions. This representation is also referred to as ambisonic representation and can be rendered binaurally or using a loudspeaker array [25]. The advantage of the ambisonic methods is that the problem of auralization is decomposed into two independent stages that can be optimized separately. From a purely practical perspective, head-tracked rendering of the ambisonic representation can be carried out straightforwardly by applying rotation operations on either the ambisonic sound field or the ambisonic HRTF representation. The application of head tracking on the rendering side in direct auralization requires computing the auralization separately for each possible head orientation of the listener, which can be a substantial computational burden.

The following subsections outline the signal processing in the auralization methods that were investigated. For the purpose of reproducibility, all implementations are provided in the Chalmers Auralization Toolbox [26, 27]. As will be discussed, the available methods were extended such that both ambisonic auralization and direct auralization can be carried out based on cubical volumetric (CV) sampling and based on spherical surface (SS) and cubical surface (CS) sampling of the virtual sound field. Fig. 2

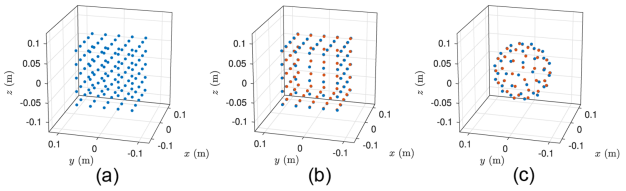


Fig. 2. Examples of the employed sampling grid types. (a) CV (125 nodes). (b) CS grid (98 double nodes). (c) SS grid (64 double nodes). With the surface grids, blue color indicates a pressure node, and orange color indicates a coincident velocity node. The diameter of the spherical grid and the edge length of the cubical grids are 140 mm.

depicts examples for all three types of sampling grids. As elaborated further below, it is sufficient to sample only the sound pressure if a volumetric grid is used. Surface grids require sampling both sound pressure and particle velocity.

While all investigated methods are formulated in frequency domain, it appears more flexible if the auralization is carried out in practice via block-wise processing of time-domain signals because the auralization does not then need to be adapted to the length of the room response. This is also how the Chalmers Auralization Toolbox performs it. The authors highlight here that this presents challenges regarding the implementation for a variety of reasons. For example, the DC part of a given spectral representation can be undefined, or time aliasing can occur, which are both challenges that do not occur in pure frequency-domain considerations. The discussion of the implementation details is beyond the scope of this article, and readers are referred to the documentation and the code of the Chalmers Auralization Toolbox.

2.1 The Ambisonic Methods

The ambisonic methods that were investigated are based on [12], which is formulated for volumetrically sampled sound pressure fields and may be considered an extension of [10]. The SH representation is obtained via inversion of a matrix of the SH basis terms. Contrary to [12], this study uses a singular value decomposition (SVD) [28] that is regularized by limiting the dynamic range of the singular values to regularize the matrix inversion instead of soft-clipping of the matrix elements followed by an unregularized pseudo inverse. SEC. 7 discusses the significance of the choice of regularization.

Volumetric sampling may be inefficient in that it may require unnecessarily many sampling points to yield a well-conditioned problem. The Kirchhoff-Helmholtz integral [29] demonstrates that the sound pressure field inside a source-free domain is uniquely described by the sound pressure $S(\vec{\mathbf{x}}, \omega)$ and normal sound pressure gradient $\frac{\partial}{\partial \vec{\mathbf{n}}} S(\vec{\mathbf{x}}, \omega)$ distributions along the simply connected surface that encloses the domain. It is proven in [16, p. 207] that a weighted sum $S(\vec{\mathbf{x}}, \omega) + \gamma \frac{\partial}{\partial \vec{\mathbf{n}}} S(\vec{\mathbf{x}}, \omega)$ uniquely defines the sound pressure inside the volume that is enclosed by the surface so long as $\mathrm{Im}\{\gamma\} \neq 0$. Choosing $\gamma = 1/(\mathrm{i}\frac{\omega}{c})$ fulfills the uniqueness criterion and is a particularly convenient choice as it assures that pressure and gradient are added with similar magnitudes. It also allows for the weighted sum to be interpreted as a virtual sensor with (far-field) cardioid directivity, which connects the formulation well to the literature on microphone arrays [15, 17]. This enables the employment of surface sampling grids, which may potentially be more efficient than volumetric ones.

The sound pressure gradient can be straightforwardly computed from the particle velocity [30], which a number of simulation methods provide straightforwardly. If the particle velocity is not available, then it can be obtained from sampling the sound pressure along two surfaces that are much closer to each other than the wavelength.

Extending [12] to SS grids with radially outward-facing cardioid sensors is straightforward given the literature on spherical microphone arrays [15]. The mathematical details are provided in the documentation of the Chalmers Auralization Toolbox. As far as the authors are aware, cardioid sensors have only been used with circular and spherical contours for performing SH sound field decomposition. This is unfortunate as some wave-based simulation methods such as finite-difference time-domain can employ Cartesian sampling schemes that would require interpolation to realize a spherical sampling grid. To bridge this gap, the method from [12] was extended to arbitrary simply connected surface grids with normally outward-facing cardioid sensors. CS grids were included in the present investigation. This solution was not available in the literature and is an original contribution by the authors. The mathematics behind it is not straightforward, and readers are referred to the documentation of the Chalmers Auralization Toolbox for details [31].

There are two main sources for systematic inaccuracies in ambisonic auralization: (1) spatial aliasing due to the discretization of the sampled sound field and (2) truncation of the order N of the SH representation of the sound field (and consequently of the HRTFs). Spatial aliasing typically increases the energy at high frequencies, and order truncation decreases or increases it depending on the incidence direction [32]. Anticipating the interaction between the two is not straightforward.

Contrary to the original formulations of ambisonic auralization [10, 12], a magnitude least-squares (MagLS) SH representation of the HRTFs is employed [33], which essentially eliminates the effect of the order truncation on the HRTFs. MagLS SH representations achieve an increase of the accuracy of the magnitude at the expense of a modification of the phase. It was demonstrated in [33] that the phase modification is not audible if it occurs at frequencies higher than 3 kHz. This condition is fulfilled for all the data that use an SH order of 6 or higher. Fig. 3 illustrates the increase of the accuracy of the magnitude exemplarily for an SH order of $N = 4$. Fig. 4 depicts numerical data for example auralizations.

The effects of spatial aliasing and order limitation in the decomposed sound field occur in the same frequency range. Their effect on the distribution of the energy over the frequency axis is hardly dependent on the incidence direction of the sound field for the grids that are considered here. Therefore, the sound field decomposition is equalized by

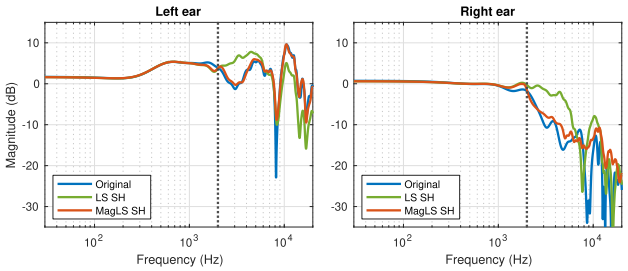


Fig. 3. The employed HRTFs (blue line) as well as their SH representations for sound incidence from the left computed from a fourth-order SH decomposition using the conventional LS fit of SH coefficients (green line) as well as using a MagLS fit (orange line). The gray dotted line indicates the frequency below which LS and above which MagLS was used to compute the SH coefficients in the case that is labeled MagLS.

computing a global minimum-phase filter that minimizes the deviation of the binaural output from the ground truth with this deviation being averaged over many sound incidence directions. This equalization filter is only effective in the frequency range where spatial aliasing is apparent. Recalling that HRTFs are defined as the acoustic ear signals due to a plane wave in free-field conditions makes sampled plane waves and HRTFs a convenient set of input/output data based on which this equalization filter can be computed.

Independent of the sampling grid, the ambisonic methods produce an SH representation of the sampled sound field. The binaural rendering of these is well established and detailed, for example, in [34].

2.2 The Direct Methods

The direct auralization methods are based on [11]. The auralization is modeled as a MIMO system. The input to the MIMO system is the sampled sound field data. The output are the left and right ear signals. The set of filters that represent the transfer path between the input and the output are computed through an LS fit on sample data, which is regularized by limiting the dynamic range of the singular values of an SVD. Sampled plane waves and HRTFs are a convenient set of input/output data for computing the LS solution. The original method was demonstrated for volumetrically sampled sound pressure. Applying it to SS and CS sampled data does not require any modifications.

Like any other method, direct auralization also suffers from spatial aliasing (cf. the numerical data depicted in Fig. 4). The LS solution can exhibit large errors in the frequency range where spatial aliasing occurs. It was found that the solution is perceptually more favorable if the LS solution is replaced with a MagLS solution in the frequency range where the spatial aliasing occurs. This implementation corresponds to what is termed end-to-end MagLS, variant 2, in [35].

3 GRID PARAMETERS

Choosing the parameters for a grid with a given number of nodes L is a compromise: a grid with large dimensions is desired for being able to resolve spatial information at low frequencies. Keeping the spatial aliasing frequency high requires a small spacing between the nodes. It was demonstrated in [32] that high SH order spatial information at low frequencies in the sound field does not reach the ear of the listener because it is suppressed by the directivity of the ears. The sampling grid can therefore be small compared to the wavelength at those low frequencies where detailed spatial information is not required. This produces a high spatial aliasing frequency, which is favorable and comes at virtually no cost. The authors verified using the approach from [32] that a diameter or edge length of 0.14 m is the smallest size that avoids compromises in the binaural output signals due to the required regularization. It was chosen for all grids. No noteworthy benefit in choosing a different grid size for any of the investigated configurations could be found.

Solutions for optimal placement of sampling points in a volume or on a surface contour for sound field decomposition have been proposed in the literature [36–38]. These have only been demonstrated in narrow frequency ranges. It is unclear how they can be extended to the entire audible frequency range that is of interest here. Therefore, only evenly sampled grids are considered in this article. This means that only certain numbers of sampling nodes are possible for CV and CS grids. The layouts for SS grids were selected from Fliege grids [39] with comparable numbers of nodes.

For ambisonic auralization, the frequency above which spatial aliasing occurs can be estimated with the $N = kR$-rule, where N is the maximum order that the array provides, $k = 2\pi f_\mathrm{a}/c$ is the wavenumber that is considered, R is the radius of the baffle, and c is the speed of sound [40]. Solving the expression for the aliasing frequency f_a yields

$$f_\mathrm{a} = \frac{N}{2\pi R} c\,. \tag{1}$$

The maximum order N that Fliege grids with L nodes on SSs provide is given by $(N+2)^2 = L$. As an example, an SS grid with $R = 0.07$ m and $L = 144$ supports $N = 10$ and produces an aliasing frequency of $f_\mathrm{a} = 7.8$ kHz. Slightly more nodes appear to be required for CSs for a given N and quite a few more for volumetric grids. The aliasing frequency f_a does not seem to differ substantially between ambisonic and direct methods.

4 COMPUTATION OF THE GROUND TRUTH

Validation of the perceptual transparency of the auralization requires comparing the auralization to the ground truth ear signals. Anechoic conditions are straightforward given that HRTFs are defined as the acoustic ear signals due to a plane wave in free-field conditions. Therefore, spatially sampled plane waves were used as the input to the auralization and the HRTFs as the ground truth in this setting. The sound pressure $S(\omega, \vec{x})$ at radian frequency ω and at location $\vec{x}$ due to a plane wave is given by $S(\omega, \vec{x}) = \mathrm{e}^{-i\vec{k}_\mathrm{pw}^T\vec{x}}$ where $\vec{k}_\mathrm{pw}^T$ is the transposed wave vector of the plane wave. The sound pressure gradient is computed accordingly.

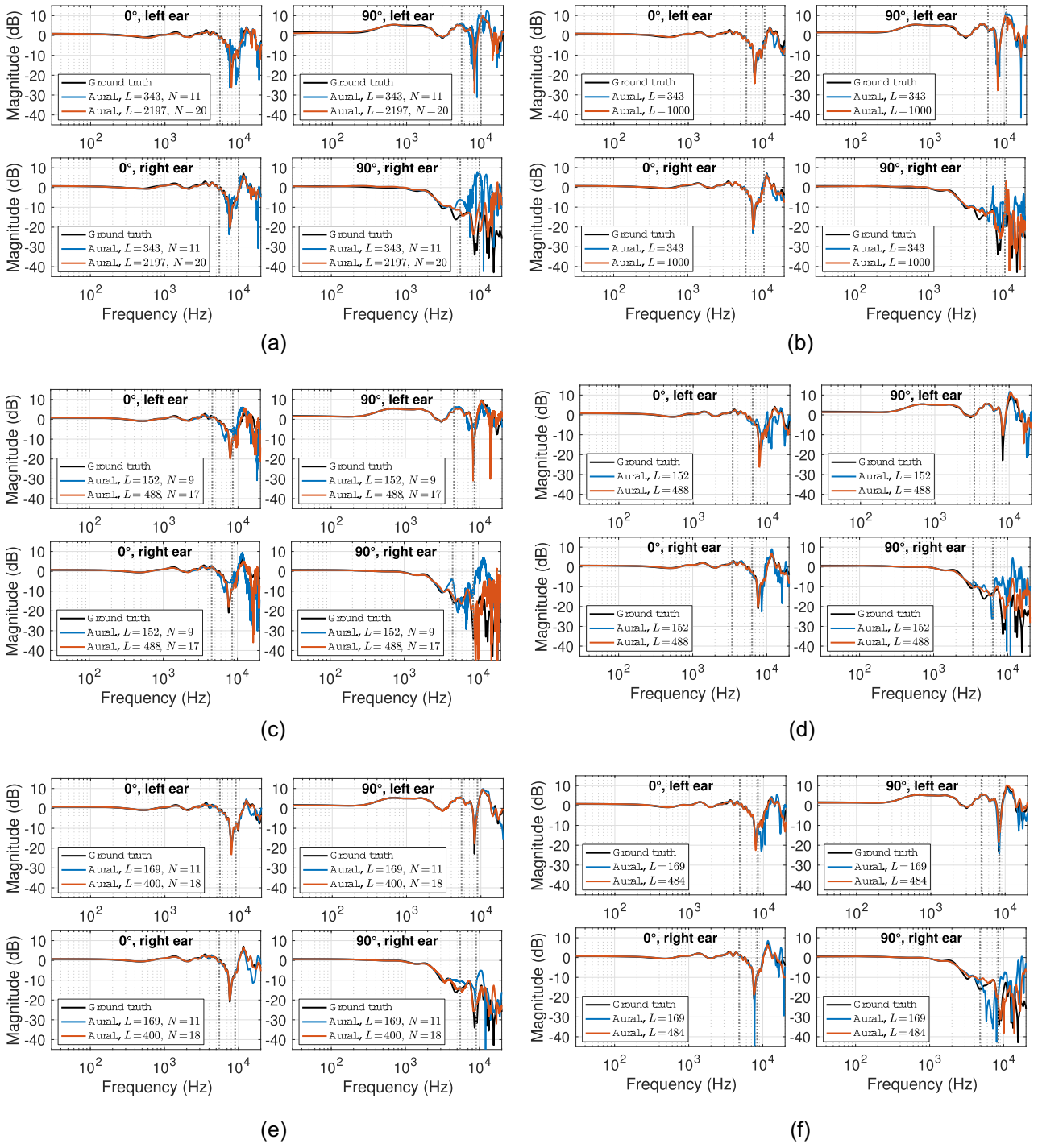


Fig. 4. Example anechoic binaural data for each combination of grid and method for a grid with a moderate number of sampling points and a high-density grid each. (a) Cubical volume grid, ambisonic auralization. (b) Cubical volume grid, direct auralization. (c) Cubical surface grid, ambisonic auralization. (d) Cubical surface grid, direct auralization. (e) Spherical surface grid, ambisonic auralization. (f) Spherical surface grid, direct auralization. The binaural output is due to a plane wave with incidence from straight ahead or from 90° to the left. The ground truth is the HRTF corresponding to the given incidence direction. The vertical dotted lines indicate above what frequency MagLS-HRTFs are employed (ambisonic methods) or the eMagLS solution is employed (direct methods). All data were computed by performing the auralization on the time-domain signals.

Reverberant conditions are less straightforward in that no wave-based simulation method is free of systematic errors, and it is not straightforward how to compute the ground truth ear signals. Simpler room acoustic simulation methods like the image source method allow for computing a ground truth, but the resulting reverberation can sound unnatural.

Therefore, the following procedure was chosen: The spatial decomposition method (SDM) [41] produces a representation of a room impulse response in terms of the pressure impulse response as well as an incidence direction for each digital sample of the pressure impulse response. A binaural representation of the room impulse response can

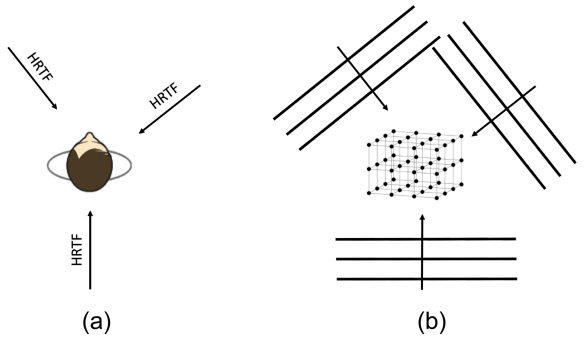


Fig. 5. Illustration of the computation of the ground truth binaural signal (a) and the sampled sound field data (b) based on SDM data as performed for the reverberant condition.

be obtained by assigning the HRTF that corresponds to the incidence direction of a given digital sample to that digital sample and superposing the weighted HRTFs for all digital samples. This binaural representation has been shown to be perceptually slightly different from the actual binaural room response. It was chosen for the present purpose because it was confirmed by different authors that the perceptual result is highly plausible [42, 43]. The input data to the auralization can be computed from the SDM data by assigning plane waves with appropriate incidence directions instead of HRTFs to each digital sample. Fig. 5 illustrates the concept. An implementation is available at [44].

5 PERCEPTUAL EXPERIMENT

The experiment was a triangular test to determine if a difference between the stimulus and the ground truth was detectable [45–47]. The subjects were presented with a graphical interface with three buttons in each trial. Two of the buttons played identical signals (ground truth or auralization), and the third one played the corresponding other signal. The assignment of the buttons as well as the order of the conditions were randomized for each participant. The subject's task was to report via a button click which of the three stimuli sounded different from the other two. They were instructed to click any of the three answer buttons randomly if they were not detecting an audible difference. The subjects could listen to the stimuli as many times as they wanted. Switching from one stimulus to another started playback of the source signal from the beginning.

A 5-s segment of a rock drum rhythm was chosen as the test signal. It is transient, which makes it suitable for investigating the auralization of reverberation, and it is one of the practically relevant signals that has relatively much energy at high frequencies, which ensures that the consequences of spatial aliasing are reflected in the binaural output with sufficient prominence.

The authors presented preliminary results from a similar validation study on the Chalmers Auralization Toolbox in [48]. That experiment used a transformed two-up/one-down staircase procedure [49] to determine the minimum required number of sampling points for making the difference between the stimulus and ground truth to be undetectable. Head tracking was employed, and the subjects were able to switch the principal direction of the sound source between straight ahead and lateral. Further experimentation with the same signals suggested that the amount of degrees of freedom that the subjects in the experiment from [48] were required to master was so high that it produced a significant number of unforced errors. Only a limited amount of data was therefore available for grids with high density of sampling points, which made it difficult to interpret the results. Furthermore, the experiment was based on the assumption that reducing the energy of spatial aliasing to a sufficient extent by increasing the number of sampling points will necessarily lead to perceptually transparent auralization. SEC. 7 discusses why the authors are doubtful that this assumption is justified.

The authors chose to present a study here that aims at validating perceptual transparency for a smaller set of auralization parameters in a more robust manner than [48] at the cost of not identifying the minimum required number of sampling points. The present experiment does not employ head tracking, and the directions of incidence of the direct sound are preselected to be 0°, –45°, and 90°. It is important to include a variety of sound incidence directions as it was confirmed in [50, 51] and other locations in the literature that lateral sound incidence is the most revealing one for detecting perceptual differences between reference and stimulus. This is presumably because the target signal at the contralateral ear exhibits comparatively low energy so that artifacts, for example due to spatial aliasing, become perceptually more prominent. The artifacts at the contralateral ear are even numerically prominent for most conditions depicted in Fig. 4.

The room acoustic conditions that were tested were anechoic and a hall with a reverb decay time of 1 s. Loudness alignment of the stimuli was not necessary as all tested methods preserve the signal level. The playback level was the same for all subjects and was similar to that of conversational speech.

The authors chose to include both low-density grids as well as high-density grids in the study. Low-density grids exhibit so few sampling points that it is expected that the subjects will be able to reliably differentiate the auralization from the ground truth. The purpose of including these grids in the study was to ensure that there are conditions for which the subjects are able to carry out their task with great confidence. Low-density grids were ommitted for the reverberant condition to minimize the experiment duration. It is not to be expected for these low-density grids that the subjects' responses to reverberant stimuli would be considerably different than to the anechoic stimuli [48].

The authors chose the high-density grids such that they exhibited so many sampling points that a saturation of the perceptual difference between the auralization and the ground truth was found in extensive informal listening. In other words, it appeared that a further increase of the number of sampling points in the high-density grids did not reduce the perceptual difference to the ground truth.

Table 1 summarizes the conditions that were tested in the experiment. The experiment comprised a total of 54 con-

Table 1. Number L of data points for the different grids for ambisonic (a) and direct (d) auralization.

	Low Density (Anechoic)	High Density (Anechoic)	High Density (Reverberant)
CV(a)	216 (7)	2,197 (20)	2,197 (20)
CV(d)	27	1,000	1,000
CS(a)	98 (7)	488 (17)	488 (17)
CS(d)	98	488	488
SS(a)	25 (3)	400 (18)	400 (18)
SS(d)	25	484	484

Note. See Fig. 2 for illustrations of the geometries. The value in parentheses states the maximum SH order N that can be extracted from the grid. Note that L for the surface grids refers to the number of pairs of data points. The cubical grids have an edge length of 0.14 m, and the spherical grids have a diameter of 0.14 m.

ditions that were tested (six anechoic conditions with low-density grids, six anechoic conditions with high-density grids, and six reverberant conditions with high-density grids; three different sound incidence directions for each of them) preceded by a training comprising eight conditions.

The HRTFs from [52] were used to compute the stimuli. Audio examples of the conditions tested in the present experiment are available in [53].

Nineteen subjects (of which 12 were men, six were women, and one preferred not to tell) between 23 and 42 years of age (median: 28 years) with self-reported unimpaired hearing participated in the experiment. The subjects were acoustics researchers, were students in the Sound and Vibration M.Sc. program, or attended a course on music technology, as well as one student without such background. The median duration of the experiment was 19 min (minimum: 10 min, maximum: 32 min).

6 RESULTS

Figs. 6–8 depict the obtained detection rates (i.e., the percentage of trials in which the odd stimulus was correctly reported, for the different tested conditions). The standardized method to evaluate the results of a triangle test is based on a binomial model. If all subjects give random answers (i.e., if the auralization is perceptually transparent), the expected detection rate is $33.\bar{3}$ %. By defining an α-level (the maximum acceptable probability of concluding that a perceptible difference exists when one does not), one can use the binomial distribution to calculate the minimum number of correct responses required to conclude that a perceivable difference exists. For the $N = 19$ subjects in this experiment, at least 11 correct detections (a detection rate of 57.9%) are needed to reject the null hypothesis of perceptual similarity at $\alpha = 0.05$ [46].

However, the absence of evidence for a perceivable difference does not imply evidence for the absence of a difference. A detection rate below the threshold does not automatically indicate perceptual transparency. To formally test for similarity, one can define a similarity threshold, p_d, which corresponds to the maximum allowable proportion of participants who are true "discriminators" (i.e., able to perceive a difference). This should not be confused with the raw detection rate, as nondiscriminators can still guess correctly by chance. To account for this, one defines a β-level, limiting the probability of a Type II error (i.e., falsely concluding similarity when a perceptible difference exists), and uses the normal approximation to the binomial distribution to calculate an upper confidence limit for p_d, as described in [46].

This study sets $p_d = 0.5$ and $\beta = 0.05$, resulting in a threshold of up to nine correct identifications (47.4%) for the stimuli to be considered sufficiently similar. This means

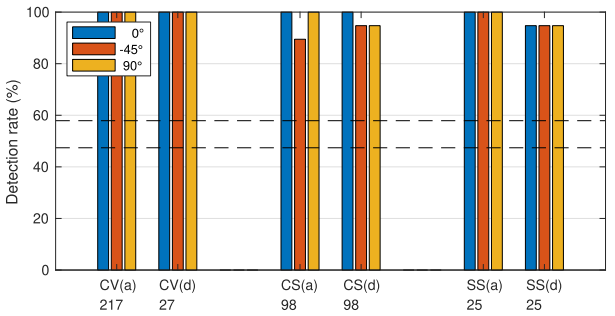


Fig. 6. Obtained detection rates for the anechoic condition for low-density grids. The dashed lines mark the maximum detection rate threshold for similarity (47.4%) and the minimum detection rate threshold for significant differences (57.9%).

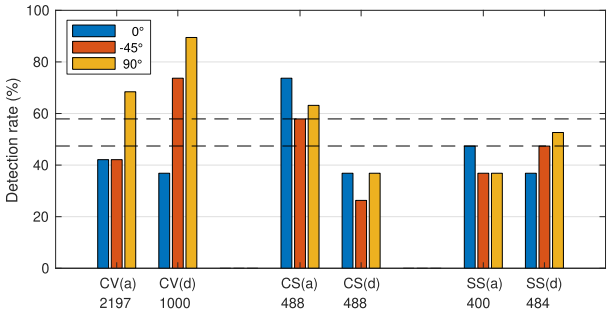


Fig. 7. Obtained detection rates for high-density grids (anechoic condition). The dashed lines mark the maximum detection rate threshold for similarity (47.4%) and the minimum detection rate threshold for significant differences (57.9%).

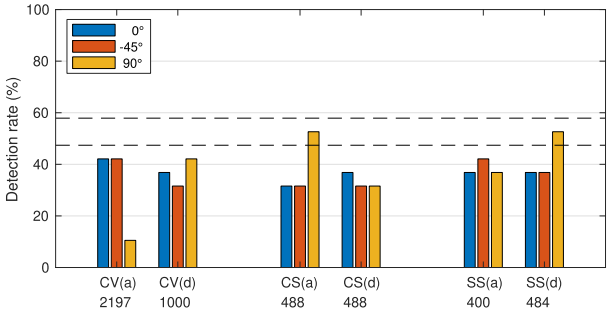


Fig. 8. Obtained detection rates for high-density grids (reverberant condition). The dashed lines mark the maximum detection rate threshold for similarity (47.4%) and the minimum detection rate threshold for significant differences (57.9%).

that even if up to 50% of the subjects can detect a difference, the stimuli are still considered perceptually similar within the chosen confidence level of 95%. While this threshold may appear generous for claiming perceptual transparency, the authors argue that the triangle test represents a highly sensitive method for difference detection. First, such a direct comparison to the ground truth is rarely possible in applied settings. Second, any practical acoustic simulation method introduces systematic uncertainties (e.g., numerical dispersion), which are likely greater than the residual deviations observed under idealized auralization conditions with high-resolution grids and anechoic conditions.

The percentage of correctly identified differences between auralization and ground truth accumulated over all subjects are depicted in Figs. 6–8, with horizontal lines indicating both the previously discussed minimum threshold for statistically significant differences (57.9%) and the maximum threshold for sufficient similarity (47.4%). The results for low-density grids under anechoic conditions are shown in Fig. 6 and confirm that participants were clearly able to identify differences independent of grid type, auralization method, and incidence angle.

For high-density grids in anechoic environments, the perceptual results shown in Fig. 7 are not as consistent. However, except for the 90° incidence direct condition, all detection rates for the SS grid are below the similarity threshold, which means perceptual transparency is assumed. While the direct auralization of the CS grid appears to be indistinguishable from the ground truth as well, the detection rate for the ambisonic auralization of the same grid exceeds the threshold for significant differences independent of incidence angle. For the high-density volumetric cube in anechoic conditions, the incidence angle appears to have a greater impact on perceivable differences, as the detection rates for 0° auralizations are well below the similarity threshold, whereas the detection rates for 90° auralizations exceed the threshold for significant differences.

Finally, the results for the high-density grids in a reverberant sound field shown in Fig. 8 confirm that, for most of the evaluated conditions, the auralizations can be considered as perceptually transparent compared to the ground truth. The only exceptions are the CS grid auralized via ambisonic and the direct auralization of the SS grid, which both slightly exceeded the similarity threshold for 90° incidence.

7 DISCUSSION

It was reported in the literature that authentic binaural reproduction of reverberation can be performed at considerably lower SH orders than the direct sound [54, 55]. This refers to cases where direct sound and reverberation were auralized separately. The targeted applications in the present case do not allow for separating direct sound and reverberation before auralization. The present experiment suggests that reverberation reduces perceived differences between auralization and ground truth somewhat. Results from [48] suggest that the properties of the direct sound appear to dominate the detectability of differences also in the reverberant case. Reverberation does not seem to have a large effect on the required signal processing for perceptual transparency. A final proof is to be presented.

The study that was presented in [55] investigated the required SH order for perceptually transparent ambisonic auralization of recordings from rigid spherical microphone arrays, which is a setup that is very similar to the SS grids in this study. The authors found that the median SH order required was approximately 20, which is in a similar order of magnitude like the SH order required for perceptual transparency in the present case. A large amount of informal listening was conducted that suggests that the results from [55] (and from [50, 51, 54] for that matter) are only partially comparable to the present ones in terms of the requirements for perceptual transparency that were determined. The reason for this is that these works used a reference signal for the comparison that was the output of a very-high-SH-order version of the rigid spherical microphone array that was being investigated. The consequence is that in [55, 50, 51, 54], the reference and the stimuli under test differed only with respect to spatial aliasing but were virtually identical below the spatial aliasing frequency.

In the present study, the ground truth was not processed by the array pipeline. It was found that this does make a noticeable difference when an experiment paradigm that is as sensitive as the triangular test is being used. The signal processing from [12, 55, 15] was tested in extensive informal listening in the present context and did not succeed in producing auralizations that sounded nearly as close to the non-array ground truth as when the present signal processing with the regularization based on the SVD is employed (recall SEC. 2). A similarly regularized SVD was already identified in [56] to provide higher numerical accuracy with respect to a variety of metrics particularly at low frequencies. The authors confirm that it is preferable also in perceptual terms and provides benefits over the entire frequency range. All stimuli that were produced with other regularization methods were clearly not perceptually transparent.

It was found that the generally accepted assumption that arrays produce sufficiently accurate binaural output signals below the spatial aliasing frequency is not strictly applicable. The authors are confident to state that in the cases where their subjects were hearing differences for the high-density grids, the contribution of spatial aliasing to the difference was small. This is supported also by the numerical data shown in Fig. 4(d): condition CS(d) 488 is perceptually transparent, but the corruption of the signal in the contralateral ear due to spatial aliasing is considerable for the 90° incidence direction.

The informal experimentation revealed that differences can even be apparent when the source signal is lowpass filtered such that no spatial aliasing is being triggered. This is the reason why the authors chose not to aim at identifying a minimum number of sampling points that guarantees perceptual transparency. Increasing the number of sampling points reduces spatial aliasing but does not reduce the possible differences to the ground truth below the spatial aliasing frequency. The authors have not been able to identify these

differences below the spatial aliasing frequency numerically, and they will direct future research toward clarifying this aspect before identifying the minimum required number of sampling points. It was found that the detectable differences below the spatial aliasing frequency occur over several octaves, which makes it unlikely that choosing different grid dimensions can mitigate this.

It appears that the exact choice of the regularization that needs to be applied with any of the tested methods is decisive. A substantial amount of trial-and-error led to the choice of frequency-dependent regularization parameters that were used in the preparation of the stimuli. It is conceivable that regularization parameters can be found for all tested high-density grids to produce perceptually transparent auralization in the strict sense.

Readers need to be reminded that the above discussion assumes comparing auralization and ground truth in a triangular test. Other settings like casual listening or potentially even A-B comparisons (without a third alternative) may not reveal the differences that are discussed above. Readers are encouraged to experience this for themselves based on the audio examples that are provided in [53].

The work presented in [18] also used virtual cardioid sensors for ambisonic auralization of sampled sound fields and confirmed high numerical accuracy of the binaural signals for an SH order of 16. Note that the authors of [18] used slightly different signal processing than the present authors, and only data up to a frequency of 10 kHz are shown. Perceptual data are not provided.

8 CONCLUSION

This work showed that the perceptually transparent auralization over the entire audible frequency range was possible with some of the tested methods and parameter choices. This is an important result as it ensures that data from any numerical simulation methods that can compute sound pressure and particle velocity at the required sampling points can be auralized faithfully. The validation paradigm employed was very sensitive so that even parameter sets that did not provide perceptual transparency in the strict sense are likely to be very useful in practice given the typical numerical uncertainties and bandwidth limitations of simulation methods.

It can be confirmed through extensive experimentation that details of the implementation of the signal processing can be decisive for whether perceptual transparency can be achieved or not. A free implementation of all investigated methods for reproducibility is provided[1]. It is not inconceivable that further improvements of the signal processing can allow for even more accurate auralization. Specifically, it was found that audible differences can remain below the spatial aliasing frequency even when carefully selected signal processing parameters are employed. Overcoming this is subject to future research.

[1] Implementations of the methods that were investigated in this article are available freely in the Chalmers Auralization Toolbox [27]. Audio examples are available at [27, 53].

It may not be convenient in many situations to design the sampling grid on which a given acoustic simulation is computed such that the simulation grid is most favorable for auralization. Interpolation needs to be applied in these cases. The perceptual implications of this are to be investigated.

The investigated methods have the potential of serving as a universal auralization methods as they can be applied to any acoustic simulation data so long as the data can be converted to a sound pressure and/or particle velocity distribution. The simulation data can originate from a variety of source including geometric acoustics simulations, wave-based simulations, and radiocity. This can make simulation data more straightforward to compare and allows for broadening the scope of round-robin tests like [3]. The code base that accompanies this article provides example projects for auralization of data from commercial room acoustic simulation software with the presented methods.

9 REFERENCES

[1] M. Vorländer, *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality* (Springer, Berlin, Germany, 2008). https://doi.org/10.1007/978-3-540-48830-9.

[2] B. N. J. Postma and B. F. G. Katz, "Perceptive and Objective Evaluation of Calibrated Room Acoustic Simulation Auralizations," *J. Acoust. Soc. Am.*, vol. 140, no. 6, pp. 4326–4337 (2016 Dec.). https://doi.org/10.1121/1.4971422.

[3] F. Brinkmann, L. Aspöck, D. Ackermann, et al., "A Round Robin for Room Acoustical Simulation and Auralization," *J. Acoust. Soc. Am.*, vol. 145, pp. 2746–2760 (2019 Apr.). https://doi.org/10.1121/1.5096178.

[4] M. Blau, A. Budnik, M. Fallahi, et al., "Toward Realistic Binaural Auralizations – Perceptual Comparison Between Measurement and Simulation-Based Auralizations and the Real Room for a Classroom Scenario," *Acta Acust.*, vol. 5, paper 8 (2021 Jan.). https://doi.org/10.1051/aacus/2020034.

[5] A. Southern, D. T. Murphy, and L. Savioja, "Spatial Encoding of Finite Difference Time Domain Acoustic Models for Auralization," *IEEE Trans. Audio Speech Lang. Process.*, vol. 20, no. 9, pp. 2420–2432 (2012 Nov.). https://doi.org/10.1109/TASL.2012.2203806.

[6] R. Mehra, L. Antani, S. Kim, and D. Manocha, "Source and Listener Directivity for Interactive Wave-Based Sound Propagation," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 4, pp. 495–503 (2014 Apr.). https://doi.org/10.1109/TVCG.2014.38.

[7] S. Bilbao, A. Politis, and B. Hamilton, "Local Time-Domain Spherical Harmonic Spatial Encoding for Wave-Based Acoustic Simulation," *IEEE Signal Process. Lett.*, vol. 26, no. 4, pp. 617–621 (2019 Apr.). https://doi.org/10.1109/LSP.2019.2902509.

[8] I. Henderson, A. Politis, and S. Bilbao, "Filter Design for Real-Time Ambisonics Encoding During Wave-Based Acoustic Simulations," in Proc. *Forum*

Acusticum, pp. 517–521 (Lyon, France) (2020 Dec.). https://doi.org/10.48465/fa.2020.0652.

[9] J. A. Hargreaves, L. R. Rendell, and Y. W. Lam, "A Framework for Auralization of Boundary Element Method Simulations Including Source and Receiver Directivity," *J. Acoust. Soc. Am.*, vol. 145, pp. 2625–2637 (2019 Apr.). https://doi.org/10.1121/1.5096171.

[10] B. Støfringsdal and P. Svensson, "Conversion of Discretely Sampled Sound Field Data to Auralization Formats," *J. Audio Eng. Soc.*, vol. 54, no. 5, pp. 380–400 (2006 May).

[11] M. A. Poletti and U. P. Svensson, "Beamforming Synthesis of Binaural Responses From Computer Simulations of Acoustic Spaces," *J. Acoust. Soc. Am.*, vol. 124, no. 1, pp. 301–315 (2008 Jul.). https://doi.org/10.1121/1.2924206.

[12] J. Sheaffer, M. van Walstijn, B. Rafaely, and K. Kowalczyk, "Binaural Reproduction of Finite Difference Simulations Using Spherical Array Processing," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 23, no. 12, pp. 2125–2135 (2015 Dec.). https://doi.org/10.1109/TASLP.2015.2468066.

[13] D. M. Murillo Gómez, J. Astley, and F. M. Fazi, "Low Frequency Interactive Auralization Based on a Plane Wave Expansion," *Appl. Sci.*, vol. 7, no. 6, paper 558 (2017 May). https://doi.org/10.3390/app7060558.

[14] E. Hulsebos, D. de Vries, and E. Bourdillat, "Improved Microphone Array Configurations for Auralization of Sound Fields by Wave-Field Synthesis," *J. Audio Eng. Soc.*, vol. 50, no. 10, pp. 779–790 (2002 Oct.).

[15] I. Balmages and B. Rafaely, "Open-Sphere Designs for Spherical Microphone Arrays," *IEEE Trans. Audio Speech Lang. Process.*, vol. 15, no. 2, pp. 727–732 (2007 Feb.). https://doi.org/10.1109/TASL.2006.881671.

[16] A. J. Burton and G. F. Miller, "The Application of Integral Equation Methods to the Numerical Solution of Some Exterior Boundary-Value Problems," *Proc. Royal Soc. London A*, vol. 323, no. 1553, pp. 201–210 (1971 Jun.).

[17] M. R. P. Thomas, "Practical Concentric Open Sphere Cardioid Microphone Array Design for Higher Order Sound Field Capture," in Proc. *IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 666–670 (Brighton, UK) (2019 May). https://doi.org/10.1109/ICASSP.2019.8683559.

[18] M. Cosnefroy, S. Guðjónsson, and F. Pind, "Physically Accurate Binaural Reproductions From Broadband Wavebased Room Acoustics Simulations, and Comparison With Measurements," in Proc. *Immersive and 3D Audio Conf. (I3DA)*, pp. 1–7 (Bologna, Italy) (2023 Sep.). https://doi.org/10.1109/I3DA57090.2023.10289312.

[19] T. Pathre, M. Hornikx, and A. Kohlrausch, "Perceptual Evaluation of Auralizations From a Wave-Based Method in a Virtual Environment," in Proc. *Euroregio/BNAM*, pp. 105–114 (Aalborg, Denmark) (2022 May).

[20] T. YoshidaT. Okuzono, and K. Sakagami, "Binaural Auralization of Room Acoustics With a Highly Scalable Wave-Based Acoustics Simulation," *Appl. Sci.*, vol. 13, no. 5, paper 2832 (2023 Feb.). https://doi.org/10.3390/app13052832.

[21] J. Merimaa and V. Pulkki, "Spatial Impulse Response Rendering I: Analysis and Synthesis," *J. Audio Eng. Soc.*, vol. 53, pp. 1115–1127 (2005 Dec.).

[22] V. Gunnarsson and M. Sternad, "Binaural Auralization of Microphone Array Room Impulse Responses Using Causal Wiener Filtering," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 29, pp. 2899–2914 (2021 Jan.). https://doi.org/10.1109/TASLP.2021.3110340.

[23] V. Pulkki, "Spatial Sound Reproduction With Directional Audio Coding," *J. Audio Eng. Soc.*, vol. 55, pp. 503–516 (2007 Jun.).

[24] A. Politis, S. Tervo, and V. Pulkki, "COMPASS: Coding and Multidirectional Parameterization of Ambisonic Sound Scenes," in Proc. *IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6802–6806 (Calgary, Canada) (2018 May). https://doi.org/10.1109/ICASSP.2018.8462608.

[25] F. Zotter and M. Frank, *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality* (Springer, Cham, Switzerland, 2019). https://doi.org/10.1007/978-3-030-17207-7.

[26] J. Ahrens, "A Software Tool for Auralization of Simulated Sound Fields," in Proc. *Institute of Acoustics*, vol. 45, Pt. 2, pp. 1–8 (2023 Sep.). http://dx.doi.org/10.25144/16010.

[27] J. Ahrens, "The Chalmers Auralization Toolbox," https://github.com/AppliedAcousticsChalmers/auralization-toolbox (accessed 31 Oct. 2024).

[28] M. X. Cohen, *Linear Algebra: Theory, Intuition, Code* (Bucharest, Romania) (sincxpress, 2021).

[29] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions* (Elsevier, Amsterdam, the Netherlands, 2005). https://doi.org/10.1016/B978-0-08-044371-3.X5000-5.

[30] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic Press, New York, NY, 1999). https://doi.org/10.1016/B978-0-12-753960-7.X5000-1.

[31] J. Ahrens, "The Chalmers Auralization Toolbox," https://github.com/AppliedAcousticsChalmers/auralization-toolbox (accessed 1 Dec. 2025).

[32] J. Ahrens, "How Small Can a Baffled Ambisonic Microphone Array Be?" in Proc. *DAGA*, pp. 1101–1104 (Hanover, Germany) (2024 Mar.).

[33] C. Schörkhuber, M. Zaunschirm, and R. Höldrich, "Binaural Rendering of Ambisonic Signals via Magnitude Least Squares," in Proc. *DAGA*, pp. 339–342 (Munich, Germany) (2018 Mar.).

[34] J. Ahrens, "Binaural Audio Rendering in the Spherical Harmonic Domain: A Summary of the Mathematics and Its Pitfalls," Technical note v. 2, Chalmers University of Technology (2022 Feb.) https://doi.org/10.48550/arXiv.2202.04393.

[35] T. Deppisch, H. Helmholz, and J. Ahrens, "End-to-End Magnitude Least Squares Binaural Rendering of Spherical Microphone Array Signals," in Proc. *Immersive and 3D Audio: From Architecture to Automotive (I3DA)*, pp. 1–7 (Bologna, Italy) (2021 Sep.). https://doi.org/10.1109/I3DA48870.2021.9610864.

[36] G. Chardon, W. Kreuzer, and M. Noisternig, "Design of Spatial Microphone Arrays for Sound Field Interpolation," *IEEE J. Sel. Top. Signal Process.*, vol. 9, vol. 5, pp. 780–790 (2015 Aug.). https://doi.org/10.1109/JSTSP.2015.2412097.

[37] S. Koyama, Chardon Gilles, and L. Daudet, "Optimizing Source and Sensor Placement for Sound Field Control: An Overview," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 28, pp. 696–714 (2020 Jan.). https://doi.org/10.1109/TASLP.2020.2964958.

[38] S. A. Verburg, F. Elvandera, T. van Waterschoot, and E. Fernandez-Grande, "Optimal Sensor Placement for the Spatial Reconstruction of Sound Fields," *EURASIP J. Audio Speech Music Process.*, vol. 2024, no. 1, pp. 41 (2024 Dec.). https://doi.org/10.1186/s13636-024-00364-4.

[39] J. Fliege and U. Maier, "The Distribution of Points on the Sphere and Corresponding Cubature Formulae," *IMA J. Numer. Anal.*, vol. 19, no. 2, pp. 317–334 (1999 Apr.). https://doi.org/10.1093/imanum/19.2.317.

[40] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, "Intrinsic Limits of Dimensionality and Richness in Random Multipath Fields," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2542–2556 (2007 Jun.). https://doi.org/10.1109/TSP.2007.893738.

[41] S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki, "Spatial Decomposition Method for Room Impulse Responses," *J. Audio Eng. Soc.*, vol. 61, no. 1/2, pp. 17–28 (2013 Jan./Feb.).

[42] N. Kaplanis, *Perception of Reverberation in Domestic and Automotive Environments*, Ph.D. thesis, Aalborg Universitet, Aalborg, Denmark (2017).

[43] J. Ahrens, "Perceptual Evaluation of Binaural Auralization of Data Obtained From the Spatial Decomposition Method," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 65–69 (New Paltz, NY) (2019 Oct.). https://doi.org/10.1109/WASPAA.2019.8937247.

[44] J. Ahrens, "Acoustic Room Responses," https://github.com/AppliedAcousticsChalmers/acoustic-room-responses (accessed 31 Oct. 2024).

[45] M. A. Amerine, R. M. Pangborn, and E. B. Roessler, *Principles of Sensory Evaluation of Food* (Academic Press, New York, NY, 1965). https://doi.org/10.1016/C2013-0-08103-0.

[46] ISO, "Sensory Analysis – Methodology – Triangle Test," *Standard 4120:2021* (2021 Jul.).

[47] N. Zacharov (Ed.), *Sensory Evaluation of Sound* (CRC Press, Boca Raton, FL, 2018). https://doi.org/10.1201/9780429429422.

[48] J. Ahrens, "Perceptually Validated Auralization With the Chalmers Auralization Toolbox," presented at the *DAS | DAGA* (Copenhagen, Denmark) (2025 Mar.).

[49] H. Levitt, "Transformed Up-Down Methods in Psychoacoustics," *J. Acoust. Soc. Am.*, vol. 49, no. 2B, pp. 467–477 (1971 Feb.). https://doi.org/10.1121/1.1912375.

[50] M. Zaunschirm, C. Schörkhuber, and R. Höldrich, "Binaural Rendering of Ambisonic Signals by Head-Related Impulse Response Time Alignment and a Diffuseness Constraint," *J. Acoust. Soc. Am.*, vol. 143, no. 6, pp. 3616–3627 (2018 Jun.). https://doi.org/10.1121/1.5040489.

[51] J. Ahrens and C. Andersson, "Perceptual Evaluation of Headphone Auralization of Rooms Captured With Spherical Microphone Arrays With Respect to Spaciousness and Timbre," *J. Acoust. Soc. Am.*, vol. 145, pp. 2783–2794 (2019 Apr.). https://doi.org/10.1121/1.5096164.

[52] B. Bernschütz, "A Spherical Far Field HRIR/HRTF Compilation of the Neumann KU 100," in Proc. *AIA/DAGA*, pp. 592–595 (Merano, Italy) (2013 Mar.).

[53] J. Ahrens, "Perceptually Transparent Binaural Auralization of Simulated Sound Fields: Audio Examples," (2025 Oct.). https://doi.org/10.5281/zenodo.17435838 (accessed 24 Oct. 2025).

[54] I. Engel, C. Henry, S. V. Amengual Garí, P. W. Robinson, and L. Picinali, "Perceptual Implications of Different Ambisonics-Based Methods for Binaural Reverberation," *J. Acoust. Soc. Am.*, vol. 149, no. 2, pp. 895–910 (2021 Feb.). https://doi.org/10.1121/10.0003437.

[55] T. Lübeck, J. M. Arend, and C. Pörschmann, "Binaural Reproduction of Dummy Head and Spherical Microphone Array Data—A Perceptual Study on the Minimum Required Spatial Resolution," *J. Acoust. Soc. Am.*, vol. 151, no. 1, pp. 467–483 (2022 Jan.). https://doi.org/10.1121/10.0009277.

[56] A. Politis and H. Gamper, "Comparing Modeled and Measurement-Based Spherical Harmonic Encoding Filters for Spherical Microphone Arrays," in Proc. *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 224–228 (New Paltz, NY) (2017 Oct.). https://doi.org/10.1109/WASPAA.2017.8170028.

THE AUTHORS



Jens Ahrens



Leon Müller

Jens Ahrens has been a professor in the Division of Applied Acoustics at Chalmers University of Technology since 2016. His primary research interest is signal processing for acoustic transducer arrays. He received a Diplom (equivalent to an M.Sc.) in Electrical Engineering/Sound Engineering jointly from the Graz University of Technology and the University of Music and Dramatic Arts, Graz, Austria, in 2005. He completed his doctoral degree (Dr.-Ing.) at the Technische Universität Berlin, Germany, in 2010. From 2011 to 2013, he was a postdoctoral researcher at Microsoft Research in Redmond, WA, and in the fall and winter terms of 2015/2016, he was a visiting scholar at the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford University, Stanford, CA. He is a member of the editorial boards of three international scientific journals on audio signal processing.

•

Leon Müller is a Ph.D. candidate at the Division of Applied Acoustics at Chalmers University of Technology working on methods for conducting listening experiments in virtual acoustic environments. He received his B.Eng. degree in audio and video engineering from the University of Applied Sciences Düsseldorf, Germany, in 2019, and his M.Sc. degree in Sound and Vibration from Chalmers University of Technology in 2021. His research focuses on virtual acoustics and human responses to sound, including behavioral, physiological, and subjective aspects. His recent work includes studies on the localization of electric vehicles and robot sounds, as well as laboratory experiments on the effects of low-level road traffic noise.