Audio Engineering Society
Convention Paper 10208
Presented at the AES 158th Convention
2025 May 22–24, Warsaw, Poland

This paper was peer-reviewed as a complete manuscript for presentation at this convention. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Ambisonic Spatial Decomposition Method with salient / diffuse separation

Lukas Gölles[1,2] and Matthias Frank[1,2]

[1]*Institute of Electronic Music and Acoustics, 8010 Graz, Austria*
[2]*University of Music and Performing Arts Graz, 8010 Graz, Austria*

Correspondence should be addressed to Lukas Gölles (goelles@iem.at)

ABSTRACT

This paper proposes a new algorithm for enhancing the spatial resolution of measured first-order Ambisonics room impulse responses (FOA RIRs). It applies a separation of the RIR into a salient stream (direct sound and reflections) and a diffuse stream to treat them differently: The salient stream is enhanced using the Ambisonic Spatial Decomposition Method (ASDM) with a single direction of arrival (DOA) per sample of the RIR, while the diffuse stream is enhanced by 4-directional (4D-)ASDM with 4 DOAs at the same time. Listening experiments comparing the new Salient/Diffuse S/D-ASDM to ASDM, 4D-ASDM, and the original FOA RIR reveal the best results for the new algorithm in both spatial clarity and absence of artifacts, especially for its Ṡ/D-ASDM variant, which keeps the DOA constant within each salient event block.

1 Introduction

Convolution with measured room impulse responses (RIR) is a typical approach of recreating the acoustical properties of existing places. Using the Ambisonics format [1] for such RIRs has the advantage of compact microphone arrays, orientation changes through simple rotation, and playback on headphones [2] and arbitrary loudspeaker arrays [3]. The application of a tetrahedral arrangement of cardioid microphone capsules is a practical solution to capture the spatial information with first-order resolution, i.e. a first-order Ambisonic RIR (FOA RIR). However, because its spatial resolution is strongly limited, it seems favorable to work with its higher-order counterpart, the higher-order Ambisonic RIR (HOA RIR). HOA [4, 5, 6] achieves a better localization of the sound events reproduced in surrounding loudspeaker systems [7, 8], especially in off-center listening positions [9, 10, 11, 12, 13]. As higher-order microphone arrays are more expensive, require more signal processing, better calibration, and have to deal with a trade-off between spatial resolution and noise boost [14, 15], creating HOA RIRs by algorithmically upmixing measured FOA RIRs seems to be a good alternative.

Multiple algorithms for upmixing FOA RIRs have been presented that are based on different assumptions of the soundfield. The Spatial Decomposition Method (SDM) [16] and its Ambisonic counterpart ASDM [17] assume that there is only a single direction-of-arrival (DOA) per time frame present in the measured impulse response. In contrast, (Higher-Order) Spatial Impulse Response Rendering (SIRR) [18, 19] calculates multiple DOAs in different frequency bands and tries to reproduce the diffuse part of the captured sound field based on an estimate. The so-called "Directional En-

hancement By The 2 + 2 Directional Signal Estimator" (2DSE2) [20] allows one to two broadband DOAs plus two diffuse directions around each sample. As our proposed algorithm is based on ASDM, we focus on reviewing further developments and challenges here. For a perceptual comparison of upmixed FOA RIRs using HO-SIRR and (4D-)ASDM, see [21].

Since its introduction in 2013, multiple publications have successfully employed (A)SDM to achieve improved spatial resolution and perceptual similarity to dummy head measurements for a wide range of measured RIRs [22, 17, 23, 24, 25]. However, (A)SDM suffers from an artifact described as graininess or roughness [26] in the later part of the upmixed RIR when convolving it with a transient sound or directly listening to it. This artifact is caused by the violation of (A)SDMs assumption that there is only a single DOA in the late part of the RIR. While some approaches tried to reduce these artifacts by synthetically adding reflections or applying widening effects [20, 27], 4D-ASDM [21] uses only information already contained in the FOA RIR by adding 3 directions/signals arranged on a tetrahedron around the detected DOA. Experiments showed that 4D-ASDM can thus effectively reduce the graininess artifact while maintaining most of the properties of ASDM. However, the authors experience while working with ASDM and 4D-ASDM revealed that ASDM sometimes provides an increased spatial clarity due to the single DOA for direct sound events and that spatial clarity might further benefit from reducing the fluctuation of the applied DOA estimate. However, in order to reduce the graininess artifact, the late, diffuse part of the RIR should be processed by 4D-ASDM.

In this paper, we propose and evaluate a new algorithm version, that is a combination of ASDM for the salient events of the RIR (direct sound and reflections) and 4D-ASDM for the diffuse residual. The idea here is to maximize spatial clarity and, at the same time, minimize graininess artifacts. First, ASDM and 4D-ASDM are briefly revisited, and the new version is introduced including the applied salient / diffuse separation [28]. Then, a listening experiment with regard to spatial clarity and artifacts compares the original FOA RIR and HOA RIRs upmixed with ASDM and 4D-ASDM to two variants of the new algorithm. Finally, our new algorithm and the experimental results are summarized in a short conclusion.

2 Towards the new algorithm

As the new upmixing algorithm for measured FOA RIRs is a combination of ASDM and 4D-ASDM, both algorithms are revisited first.

2.1 ASDM

As an input, the Ambisonic Spatial Decomposition Method (ASDM) [22, 17] uses a 4-channel FOA RIR in B-format $\boldsymbol{h}(t) = \begin{bmatrix} w(t) & x(t) & y(t) & z(t) \end{bmatrix}^\intercal$, cf. Figure 1 (left). A single direction is assumed for each sample of the RIR, for which the pseudo-intensity vector is calculated to estimate the direction of arrival (DOA),

$$\boldsymbol{\theta}_{\mathrm{DOA}}(t) = \mathscr{F}_{\mathrm{smth}} \left\{ \mathscr{F}_{\mathrm{bp}}\{w(t)\}\, \mathscr{F}_{\mathrm{bp}} \left\{ \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} \right\} \right\}, \tag{1}$$

where $\mathscr{F}_{\mathrm{bp}}\{\cdot\}$ denotes a zero-phase bandpass filter that is applied to reduce effects of low-frequency disturbance and spatial aliasing of the microphone array. To prevent the DOA from jumping rapidly, its fluctuations are smoothed by a median filter $\mathscr{F}_{\mathrm{smth}}$.

The spatially enhanced HOA RIR $\tilde{\boldsymbol{h}}(t)$ is achieved by re-encoding the omnidirectional component $w(t)$ in an arbitrarily high order N,

$$\tilde{\boldsymbol{h}}(t) = \boldsymbol{y}_N(\boldsymbol{\theta}_{\mathrm{DOA}})\, w(t)\,, \tag{2}$$

where $\boldsymbol{y}_N(\boldsymbol{\theta}_{\mathrm{DOA}}) = \begin{bmatrix} Y_0^0(\boldsymbol{\theta}_{\mathrm{DOA}}) & \dots & Y_N^N(\boldsymbol{\theta}_{\mathrm{DOA}}) \end{bmatrix}^\intercal$ contains the spherical harmonics Y up to order N evaluated at the calculated direction of arrival $\boldsymbol{\theta}_{\mathrm{DOA}}$. Finally, a spectral decay correction equalizes the third-octave energetic envelope within each spherical harmonic order to match the original FOA RIR and to suppress an increase of the reverberation time towards high frequencies [17, 21].

2.2 4D-ASDM

The Four-Directional Ambisonic Spatial Decomposition Method (4D-ASDM) [21] uses the same DOA estimation as ASDM, cf. Equation 1. However, it adds three directions that form a tetrahedron with the estimated DOA, cf. Figure 1 (right). This is done by rotating a tetrahedron prototype

$$\boldsymbol{\Theta}_0 = \begin{bmatrix} 0 & \sqrt{\frac{8}{9}} & -\sqrt{\frac{2}{9}} & -\sqrt{\frac{2}{9}} \\ 0 & 0 & \sqrt{\frac{2}{3}} & -\sqrt{\frac{2}{3}} \\ 1 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix} \tag{3}$$

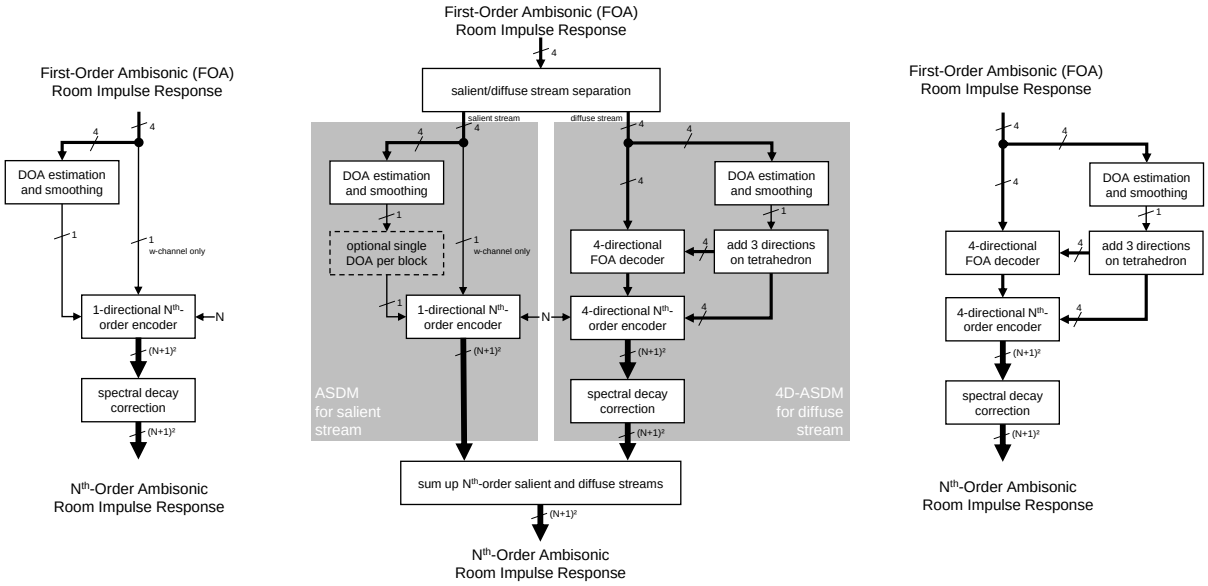


Fig. 1: Signal flow of ASDM (left), 4D-ASDM (right), and S/D-ASDM with salient/diffuse separation (center).

according to the azimuth and elevation angles φ and ϑ of the estimated DOA

$$\boldsymbol{\Theta} = \begin{bmatrix} \cos(\varphi)\cos(\vartheta) & -\sin(\varphi) & \cos(\varphi)\sin(\vartheta) \\ \sin(\varphi)\cos(\vartheta) & \cos(\varphi) & \sin(\varphi)\sin(\vartheta) \\ -\sin(\vartheta) & 0 & \cos(\vartheta) \end{bmatrix} \boldsymbol{\Theta}_0 \,. \tag{4}$$

A sampling decoder extracts the signals of four virtual microphones positioned on the edges of the tetrahedron,

$$\boldsymbol{s}(t) = \boldsymbol{Y}_1^{\mathsf{T}} \boldsymbol{h}(t) \tag{5}$$

where $\boldsymbol{Y}_1$ contains the spherical harmonics Y up to first order evaluated at the directions of the tetrahedron $\boldsymbol{\Theta}$. The spatially enhanced HOA RIR $\tilde{\boldsymbol{h}}(t)$ is achieved by re-encoding the four signals $\boldsymbol{s}(t)$ in an arbitrarily high order N,

$$\tilde{\boldsymbol{h}}(t) = \boldsymbol{y}_N(\boldsymbol{\Theta})\, \boldsymbol{s}(t)\,, \tag{6}$$

where $\boldsymbol{y}_N(\boldsymbol{\Theta})$ contains the spherical harmonics Y up to order N evaluated at the directions of the tetrahedron $\boldsymbol{\Theta}$. As performed for ASDM, a spectral decay correction is necessary as a final step.

2.3 S/D-ASDM

ASDM with salient/diffuse separation (S/D-ASDM) first splits the FOA RIR into a salient stream (direct sound and discrete reflections) $\boldsymbol{h}_{\mathrm{sal}}(t)$ and a diffuse stream $\boldsymbol{h}_{\mathrm{dif}}(t)$, cf. Figure 1 (center). This is done using an algorithm by Deppisch et. al [28] that is not specifically designed for FOA RIRs, but for multi-channel RIRs in general. It performs a generalized singular value decomposition (GSVD) and uses the singular values to decide whether a block contains a salient event or belongs to the diffuse stream. The analysis is performed backwards as it is assumed that only diffuse components are present in the late part of the RIR. In contrast to algorithms that separate the impulse response into two strictly separate streams, this algorithm keeps a diffuse component even if reflections are detected. Figure 2 shows the results of the salient/diffuse separation for one of the RIRs used in the experiment, cf. small church in Table 1.

Directional enhancement of the diffuse stream $\boldsymbol{h}_{\mathrm{dif}}$ is performed using 4D-ASDM. For the salient stream $\boldsymbol{h}_{\mathrm{sal}}$, ASDM with a single DOA is applied, however without spectral decay correction because the increased spectral brightness does only occur in the late, diffuse part of the enhanced HOA RIRs. The salient stream is divided into blocks containing a single event each. Blocks with too few samples are omitted. As an example, Figure 3 shows the azimuth and elevation of the DOA in the direct sound block of the RIR of an anechoic chamber, cf. Table 1. It is noticeable that in the part with high

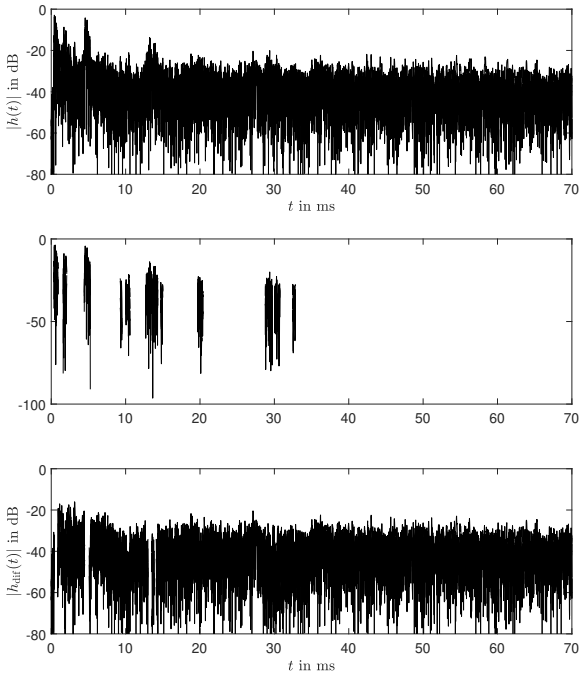


Fig. 2: RIR of the small church (top) separated in salient part (center) and diffuse part (bottom).

signal amplitude $w(t)$, the smoothed DOA still varies around 20° in azimuth and 5° in elevation.
To prevent the DOA of direct sound and salient reflections from moving around, a single, average DOA is calculated for each block. Assuming a block length of L, the averaging of the DOA is weighted by the energy of the omnidirectional signal,

$$\overline{\boldsymbol{\theta}}_{\mathrm{DOA}} = \frac{\sum_L \boldsymbol{\theta}_{\mathrm{DOA}}(t)\,|w_{\mathrm{sal}}(t)|}{\sum_L |w_{\mathrm{sal}}(t)|}\,. \tag{7}$$

This is similar to the segmentation approach in [29], however with non-continuous segments. The algorithm variant with a single DOA for each salient event block is abbreviated as Ṡ/D-ASDM, while the algorithm with the smoothed, fluctuating DOA is abbreviated as S/D-ASDM.
Finally, the enhanced HOA RIR is the sum of both parts,

$$\tilde{\boldsymbol{h}}(t) = \tilde{\boldsymbol{h}}_{\mathrm{sal}}(t) + \tilde{\boldsymbol{h}}_{\mathrm{dif}}(t)\,. \tag{8}$$

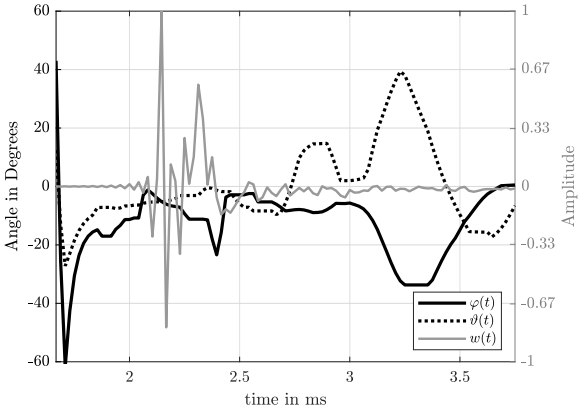


Fig. 3: Omnidirectional channel $w(t)$ of RIR in the anechoic chamber, cf. Table 1, showing azimuth φ and elevation ϑ of the smoothed DOA.

3 Listening Experiment

A listening test compares the new S/D-ASDM and Ṡ/D-ASDM algorithms to ASDM, 4D-ASDM, and the original FOA RIR in terms of spatial clarity and absence of artifacts.

3.1 Setup and Conditions

The listening experiment used measured FOA RIRs with a sampling rate of 48 kHz from three different rooms, cf. Table 1. The directional resolution of these measured FOA RIRs was enhanced to 5th order with ASDM, 4D-ASDM, S/D-ASDM, and Ṡ/D-ASDM. For all enhancement algorithms, the bandpass for the DOA estimation was set between 200 Hz and 2 kHz and the median filter length was 10 samples, as in [21]. The salient/diffuse separation was performed using a block-size of 64 samples (1.3 ms) and a hop size of 16 samples. The residual estimate had a length of 5 ms and the thresholds were calculated with a GSV averaging length of 16 blocks and using $\kappa = 5.9$ (anechoic chamber) and $\kappa = 3.5$ (studio and small chamber). In the salient stream, blocks with a length of less than 10 samples were omitted.

The experiment was carried out in the control room at IEM Petersgasse, a 5.8 m × 4.3 m × 3.2 m with a reverberation time $T_{30} \approx 0.25$ s, cf. Figure 4. For playback, we used a 7.0.4 arrangement of Genelec 8020s with a radius of 1.3 m (the loudspeakers placed on the horizon are shown in black, and the elevated loudspeakers (45°)

Room	Size	T_{60}	Microphone	Distance	Location
Anechoic chamber	65.25 m^3	-	ST450	1.3 m	IEM, Graz, Austria
Studio	141.8 m^3	0.25 s	ST450	2.0 m	IEM, Graz, Austria
Small church	2,600 m^3	1.4 s	SPS422B	11.5 m	St. Andrew's Church, Lyddington, UK[1]

Table 1: FOA RIRs in the experiment with room size, reverberation time, microphone type, microphone/loudspeaker distance, and measurement location.

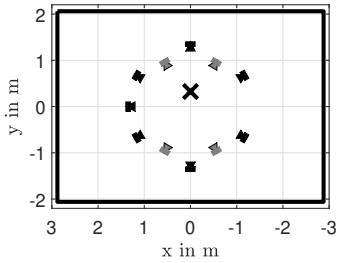


Fig. 4: Loudspeaker setup (black: 0° elevation; grey: 45° elevation) at the IEM Petersgasse with listening position marked as cross.

are shown in gray). Decoding to the loudspeaker arrangement was performed using the AllRAD approach [3] without any additional weighting. The imaginary loudspeakers at nadir and zenith were set to a gain of zero. We decided to use a listening position slightly off-center since it helps to hear differences between the algorithms and is more representative for a larger number of listeners.

The first part of the experiment evaluated the spatial clarity and employed the first sentence of EBU's male English speech reference recording [30] convolved with the RIRs of three different rooms, cf. Table 1. In addition to two conventional rooms, a studio and a small church, we added an anechoic chamber to investigate a scenario which is most sensitive to the direct sound components. Spatial clarity was defined as the ease of localization of individual sources [31]. In addition, listeners were asked to incorporate the compactness of the different frequency components in the direct sound. After the three trials on spatial clarity, a second part evaluated the absence of artifacts, in which listeners were directly listening to the RIRs. Since the artifacts tend to affect the late reverberation, only the studio and the small church were tested, but not the anechoic chamber, resulting in two trials. Before the listening experiment started, we clarified the terms spatial clarity and artifacts and performed demonstrations.

Each trial consisted of five conditions: HOA RIRs created by ASDM, 4D-ASDM, S/D-ASDM, and Ṡ/D-ASDM, as well as the original FOA RIR. The latter served as a lower anchor for spatial clarity and as a reference for absence of artifacts. The comparison of the five conditions used multi-stimulus paradigm, similar to MUSHRA [32], with looped playback and cross-fades between the conditions.

3.2 Results

In total, 20 listeners (5 female) with an average age of 31 years participate in the experiment. It took them 11 min on average to complete the experiment.

The results of the first part regarding spatial clarity show that the original FOA RIR was rated worst due to the poor spatial resolution of FOA, cf. Figure 5. In all three rooms, Ṡ/D-ASDM was rated best. In most rooms, 4D-ASDM achieved the worst results of all enhancement algorithms.

A signed-rank test with Bonferroni-Holm correction indicates a significant difference of all upmixed HOA RIRs from the original FOA RIRs in the anechoic chamber ($p < 0.0008$). The results of the anechoic chamber exhibit a significant difference when comparing Ṡ/D-ASDM with all other enhancement algorithms ($p < 0.006$), as the fluctuating DOA of the direct sound becomes most apparent here, cf. Figure 3. Considering the impact effect size [33], any pairwise comparison between the algorithms indicates great effects ($d > 1$), except for a small effect ($d = 0.39$) between ASDM and 4D-ASDM, cf. Table 2.

In the studio, the signed-rank test reveals significant differences between ASDM and Ṡ/D-ASDM, and between 4D-ASDM and Ṡ/D-ASDM ($p < 0.0007$), but only a weakly significant difference between S/D-ASDM and Ṡ/D-ASDM ($p = 0.094$). The results show the advantage of separating the salient and diffuse components as the difference between S/D-ASDM and ASDM or 4D-ASDM is significant ($p < 0.004$).

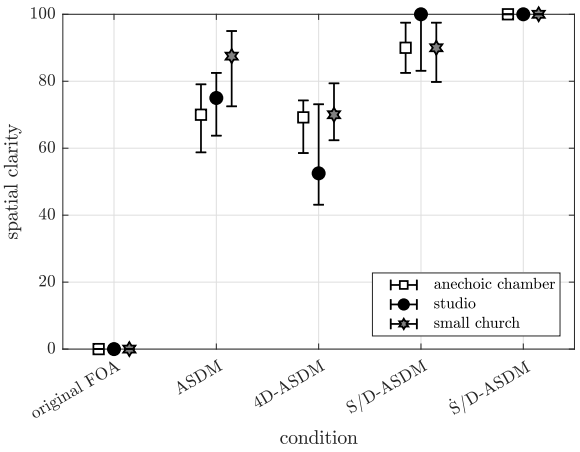


Fig. 5: Medians and corresponding 95% confidence intervals for spatial clarity.

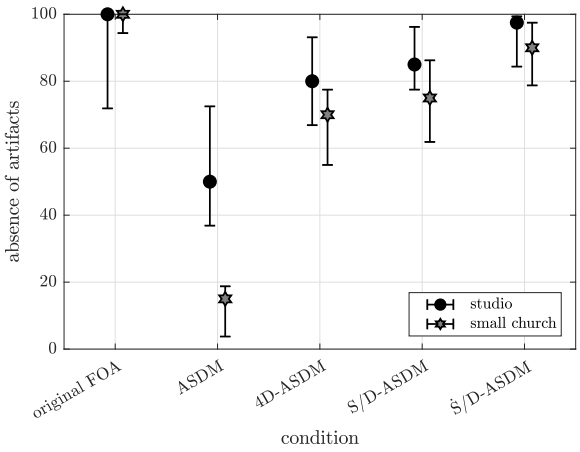


Fig. 6: Medians and corresponding 95% confidence intervals for absence of artifacts.

In the studio, large effects can be found for pairwise comparison between most algorithms ($d > 1$), except between ASDM and 4D-ASDM (medium effect, $d = 0.79$) and between S/D-ASDM and Ṡ/D-ASDM (very small effect, $d = 0.1$).
When listening to the reverberant small church, Ṡ/D-ASDM achieves the best results with significant differences to all the other conditions ($p < 0.02$). There is only a weakly significant advantage of ASDM over 4D-ASDM ($p = 0.07$) and no significant difference between ASDM and S/D-ASDM ($p = 0.59$). This can be attributed to the fact that those two algorithms differ only in their strategy of processing the diffuse stream, which is not as relevant for spatial clarity. In the most reverberant scenario, large effects are only achieved when comparing 4D-ASDM and Ṡ/D-ASDM ($d = 2.44$), for all other algorithm pairings, small to medium effects are indicated ($d < 0.69$).

In the second experimental part regarding the absence of artifacts, listeners found the original FOA RIR to have the least artifacts, however there is no significant difference to Ṡ/D-ASDM in the studio ($p = 1.3$) and the small church $p = 0.11$. ASDM is rated worst in both rooms ($p < 0.003$) and there is no significant difference between 4D-ASDM and S/D-ASDM ($p > 0.11$). Regarding the effect sizes, cf. Table 3, there are only medium effects between ASDM and 4D-ASDM for the studio ($d = 0.647$), but large effects for the comparisons between ASDM and all other algorithms (studio: $d > 0.9$, church: $d > 2.9$). Between 4D-ASDM and S/D-ASDM, there are medium effects in both rooms ($d = 0.5$) and little effects when comparing 4D-ASDM and Ṡ/D-ASDM ($d < 0.4$).

4 Conclusion

This contribution presented the S/D-ASDM algorithm for enhancing the spatial resolution of measured first-order Ambisonic room impulse responses (FOA RIR) to higher orders. It first applies a separation of the FOA RIR into salient (direct sound and reflections) and diffuse stream and treats them differently. The salient stream is enhanced using the Ambisonic Spatial Decomposition Method (ASDM) with a single direction of arrival (DOA) per sample of the RIR to maximize spatial clarity. The diffuse stream is enhanced by 4-directional (4D-)ASDM with 4 DOAs at the same time to minimize graininess artifacts.
A listening experiment reveals that the spatial clarity of S/D-ASDM is indeed comparable to ASDM, while it matches 4D-ASDM in avoiding artifacts. Moreover, its Ṡ/D-ASDM variant that minimizes directional fluctuation in the salient stream by encoding only a single DOA per salient event, could even outperform all included enhancement algorithms.
Future research should explore the applicability of S/D-ASDM to studio environments considering multichannel loudspeaker layouts and phantom sources. Moreover, the perceptual effect of lowering the enhancement order of the diffuse stream should be investigated, as lower orders would strongly reduce the computational load in real-time convolution.

algorithms	anechoic chamber	studio	small church
ASDM ǀ 4D-ASDM	0.391	0.794	0.487
ASDM ǀ S/D-ASDM	1.011	1.057	0.136
ASDM ǀ Ṡ/D-ASDM	2.878	2.107	0.623
4D-ASDM ǀ S/D-ASDM	1.019	2.250	0.687
4D-ASDM ǀ Ṡ/D-ASDM	2.808	3.118	2.44
S/D-ASDM ǀ Ṡ/D-ASDM	1.261	0.1	0.501

Table 2: Impact effect sizes for pairwise comparisons for the experimental results regarding spatial clarity.

algorithms	studio	small church
ASDM ǀ 4D-ASDM	0.647	2.910
ASDM ǀ S/D-ASDM	0.949	3.446
ASDM ǀ Ṡ/D-ASDM	1.680	4.274
4D-ASDM ǀ S/D-ASDM	0.164	0.210
4D-ASDM ǀ Ṡ/D-ASDM	0.538	0.570
S/D-ASDM ǀ Ṡ/D-ASDM	0.325	0.408

Table 3: Impact effect sizes for pairwise comparisons for the experimental results regarding the absence of artifacts.

5 Acknowledgments

We thank all listeners for their participation in the experiment.

6 Data Availability

To support reproducible research, the results of the listening experiments are available under [34].

References

[1] Zotter, F. and Frank, M., *Ambisonics - A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, Springer, 2019.

[2] Schörkhuber, C., Zaunschirm, M., and Höldrich, R., "Binaural Rendering of Ambisonic Signals via Magnitude Least Squares," in *Fortschritte der Akustik - DAGA*, Munich, 2018.

[3] Zotter, F. and Frank, M., "All-Round Ambisonic Panning and Decoding," *J. Audio Eng. Soc*, 60(10), pp. 807–820, 2012.

[4] Malham, D. G. and Myatt, A., "3D Sound Spatialization using Ambisonic Techniques," *Computer Music Journal*, 19(4), pp. 58–70, 1995, doi: http://dx.doi.org/10.2307/3680991.

[5] Poletti, M. A., "The Design of Encoding Functions for Stereophonic and Polyphonic Sound Systems," *J. Audio Eng. Soc.*, 44(11), pp. 948–963, 1996.

[6] Daniel, J., Rault, J.-B., and Polack, J.-D., "Acoustic Properties and Perceptive Implications of Stereophonic Phenomena," in *6th International Conference of the Audio Engineering Society: Spatial Sound Reproduction*, 1999.

[7] Benjamin, E., Heller, A., and Lee, R., "Localization in horizontal-only ambisonic systems," in *121th Convention of the Audio Engineering Society*, Audio Engineering Society, 2006.

[8] Braun, S. and Frank, M., "Localization of 3D ambisonic recordings and ambisonic virtual sources," in *1st International Conference on Spatial Audio,(Detmold)*, 2011.

[9] Bates, E., Boland, F., Furlong, D., and Kearney, G., "A comparative study of the performance of spatialization techniques for a distributed audience in a concert hall environment," in *31st International Conference of the Audio Engineering Society: New Directions in High Resolution Audio*, Audio Engineering Society, 2007.

[10] Frank, M., Zotter, F., and Sontacchi, A., "Localization experiments using different 2D Ambison-

ics decoders," in *25th Tonmeistertagung-VDT International Convention*, 2008.

[11] Bertet, S., Daniel, J., Parizet, E., and Warusfel, O., "Investigation on localisation accuracy for first and higher order ambisonics reproduced sound sources," *Acta Acustica united with Acustica*, 99(4), pp. 642–657, 2013, doi: 10.3813/AAA.918643.

[12] Stitt, P., Bertet, S., and van Walstijn, M., "Off-centre localisation performance of ambisonics and HOA for large and small loudspeaker array radii," *Acta Acustica united with Acustica*, 100(5), pp. 937–944, 2014, doi: 10.3813/AAA.918773.

[13] Frank, M. and Zotter, F., "Exploring the Perceptual Sweet Area in Ambisonics," in *142th Convention of the Audio Engineering Society*, 2017.

[14] Daniel, J. and Moreau, S., "Further study of sound field coding with higher order ambisonics," in *116th Convention of the Audio Engineering Society*, 2004.

[15] Rafaely, B., *Fundamentals of spherical array processing*, volume 16, Springer, 2018, ISBN 978-3-662-45664-4.

[16] Tervo, S., Pätynen, J., Kuusinen, A., and Lokki, T., "Spatial Decomposition Method for Room Impulse Responses," *J. Audio Eng. Soc.*, 61(1/2), pp. 17–28, 2013.

[17] Zaunschirm, M., Frank, M., and Zotter, F., "BRIR synthesis using first-order microphone arrays," *144th Convention of the Audio Engineering Society*, 2018.

[18] Pulkki, V., Merimaa, J., and Lokki, T., "Reproduction of Reverberation with Spatial Impulse Response Rendering," *J. Audio Eng. Soc.*, 61(1/2), pp. 17–28, 2004.

[19] McCormack, L., Pulkki, V., Politis, A., Scheuregger, O., and Marschall, M., "Higher-Order Spatial Impulse Response Rendering: Investigating the Perceived Effects of Spherical Order, Dedicated Diffuse Rendering, and Frequency Resolution," *J. Audio Eng. Soc.*, 68(5), pp. 338–354, 2020, doi: 10.17743/jaes.2020.0026.

[20] Gölles, L. and Zotter, F., "Directional Enhancement of First-Order Ambisonic Room Impulse Responses By The 2+2 Directional Signal Estimator," in *15th International Conference on Audio Mostly*, pp. 38–45, 2020, doi: 10.1145/3411109.3411131.

[21] Hoffbauer, E. and Frank, M., "Four-Directional Ambisonic Spatial Decomposition Method With Reduced Temporal Artifacts," *Journal of the Audio Engineering Society*, 70(20), pp. 1002–1014, 2022, doi: 10.17743/jaes.2022.0039.

[22] Frank, M. and Zotter, F., "Spatial impression and directional resolution in the reproduction of reverberation," *Fortschritte der Akustik - DEGA*, 2016.

[23] Ahrens, J., "Perceptual evaluation of binaural auralization of data obtained from the spatial decomposition method," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2019, doi: 10.1109/WASPAA.2019.8937247.

[24] Zaunschirm, M., Frank, M., and Zotter, F., "Binaural Rendering with Measured Room Responses: First-Order Ambisonic Microphone vs. Dummy Head," *Applied Sciences*, 10(5), 2020, ISSN 2076-3417, doi: 10.3390/app10051631.

[25] Pawlak, A., Lee, H., Mäkivirta, A., and Lund, T., "Subjective evaluation of spatial analysis and synthesis methods using different microphone arrays," in *International Conference on Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, pp. 1–7, 2021, doi: 10.1109/I3DA48870.2021.9610905.

[26] Meyer-Kahlen, N., Schlecht, S. J., and Lokki, T., "Perceptual roughness of spatially assigned sparse noise for rendering reverberation," *J. Acoust. Soc. Am.*, 150(5), pp. 3521–3531, 2021, doi: 10.1121/10.0007048.

[27] Amengual Garí, S. V., Arend, J. M., Calamia, P. T., and Robinson, P. W., "Optimizations of the Spatial Decomposition Method for Binaural Reproduction," *J. Audio Eng. Soc.*, 68(12), pp. 959–976, 2021, doi: 10.17743/jaes.2020.0063.

[28] Deppisch, T., Garí, S. V. A., Calamia, P., and Ahrens, J., "Direct and Residual Subspace Decomposition of Spatial Room Impulse Responses," *IEEE/ACM Transactions on Audio, Speech, and*

Language Processing, 31, pp. 927–942, 2023, doi: 10.1109/TASLP.2023.3240657.

[29] Pawlak, A. and Lee, H., "Spatial segmentation of an early part of the room impulse response into prominent sound events," in *Proceedings of the 24th International Congress on Acoustics*, p. 8, ICA, 2022.

[30] EBU, "EBU SQAM CD: Sound Quality Assessment Material recordings for subjective tests," 2008, EBU SQAM CD: Sound Quality Assessment Material recordings for subjective tests.

[31] Francombe, J., Brookes, T., and Mason, R., "Evaluation of Spatial Audio Reproduction Methods (part 1): Elicitation of Perceptual Differences," *J. Audio Eng. Soc.*, 65(3), pp. 198–211, 2017, doi: https://doi.org/10.17743/jaes.2016.0070.

[32] Union, I. T., "RECOMMENDATION ITU-R BS.1534-1 - Method for the subjective assessment of intermediate quality level of coding systems," *International Telecommunications Union, Geneva*, 2001.

[33] Lötsch, J. and Ultsch, A., "A non-parametric effect-size measure capturing changes in central tendency and data distribution shape," *PLoS ONE*, 15, 2020, doi: 10.1371/journal.pone.0239623.

[34] Lukas Gölles, h., "Supplementary Data for 1 to 4 - Directional Ambisonic Spatial Decomposition Method," https://anonymous.4open.science/r/Ambisonic-Spatial-Decomposition-Method-with-salient-diffuse-separation-A330/, 2025.