forum acusticum 2026
8-12 September · Graz, Austria

CHARACTERISATION OF SWEEPING ECHO INTENSITY PATTERNS IN SHOEBOX-SHAPED ROOMS: A COMPUTATIONAL STUDY

Andrea Andrijašević

Department of Electrical Engineering, University of Applied Sciences of Rijeka, Croatia

This paper is a revised version of the authors' submitted manuscript that was peer-reviewed by at least two anonymous reviewers.

Abstract

Sweeping echoes are an acoustical phenomenon present in impulse responses of shoebox-shaped rooms generated with either geometrical or wave-based room acoustics modelling software. They are also present in audio recordings convolved with the responses, introducing a distinctive change in pitch that contradicts the usual room reverberance characteristics. Based on an analytical expression derived for the rigid walls image source solution of the wave equation, it has recently been hypothesised, and supported with a modest amount of simulation data, that room setups with a higher degree of regularity lead to stronger sweeping echoes. The main objective of this study was to further test this hypothesis in shoebox-shaped rooms of various levels of regularity in room dimensions, for a number of source-receiver positions. To this end, sweeping echoes present in impulse responses obtained with the image source method were evaluated using the sweeping echoes intensity (SEI) measure. The results, based on more than 140 million of impulse responses, the corresponding SEI values, and an informal non-blind listening test, indicate that the sweeping echoes are present and audible at various source-receiver positions not only in the shoebox-shaped rooms with regular parameters but in general, i.e. irregular shoebox geometries, though to a lesser extent.

Keywords: *sweeping echo, image source method, open-source software, room acoustics modelling, shoebox-shaped room*

1. Introduction

Sweeping echoes are an acoustical phenomenon present in real-world rooms that closely resemble the ideal shoebox, i.e. rectangular cuboid, shape, such as a squash court [1]. Moreover, they can also be detected in audio recordings convolved with impulse responses of shoebox-shaped rooms generated with either geometrical or wave-based room acoustics modelling software [1], [2]. As this phenomenon introduces a distinctive linear increase in pitch of the source sound signal at the receiver position [1], contradicting the usual room reverberance characteristics, it is not very desirable in practice in the software outputs.

An analytical expression derived in [1] revealed that sweeping echoes are caused by a progressive and monotonic convergence of the time delays of specular sound reflections. Based on that expression, derived for the rigid walls image source solution of the wave equation [3], it was also hypothesised that room setups with a higher degree of regularity, namely in the values of room dimensions, and source and receiver positions, lead to stronger sweeping echoes. Experimental results obtained on a set of 30 impulse responses generated with the image source method (ISM) for a cube-shaped room, for one source position, supported the hypothesis. In [1], the authors also proposed a modification to the image source method, named randomised image method (RIM), that enables either reduction or complete elimination of sweeping echoes from the impulse responses while introducing only minor changes to the modal response. This was first accomplished by randomising the positions of image sources on the lines connecting them with the receiver [1] and, later, within small cube-shaped volumes [4].

Even though the randomisation of the positions of image sources introduces only small computational overhead when compared with the standard method, the results presented in [5], exploring the current approaches used for sweeping echoes management in open-source software, revealed that only 3 out of 17 (18 %) analysed software packages had some form of randomisation strategy implemented in the source code. These 17 packages present a subset of the 54 open-source software packages listed in [6],

Corresponding author: *andrea.andrijasevic@veleri.hr*

accompanying the study by Hornikx et al. [7], obtained by setting the following three filtering criteria:

- method: ISM or hybrid (ISM + another geometrical or wave-based method),
- main programming language: Matlab, Python, or C family,
- active repository (GitHub or SourceForge).

Given the modest number of software packages that can manage sweeping echoes currently, SEI – a measure for fast objective evaluation of sweeping echo intensity was introduced in [5]. Results for a cube-shaped room and Allen and Berkley's room obtained for a 2D grid of receiver positions of predetermined height and one source position, presented in [5], indicated that SEI can successfully identify problematic impulse responses.

Due to its low computational cost, when compared with the sweeping spectrum flatness (SSF) measure proposed in [1], it may serve as a complementary addition to the existing open-source modelling software by providing fast assessment of SEI patterns in shoebox-shaped rooms, thereby enabling informed avoidance or flagging of receiver positions with strong sweeping echoes. This synergistic approach can be of particular use in machine learning-based applications such as speech or speaker recognition, sound source localisation, speech source separation, as well as speech enhancement and dereverberation.

Building on the aforementioned research results, this study aims to:

- examine the intensity of sweeping echoes for a wide range of source and receiver positions commonly present in real-world scenarios,
- analyse the influence of room dimensions regularity,
- analyse the influence of the source position,
- determine the percentage of impulse responses for which the sweeping echoes are audible,
- examine the changes in the SEI values occurring when the randomisation of image sources is introduced.

To this end, the remainder of the paper is structured as follows. In Section 2, the details of the room geometries, of the source and receiver positions, as well as of the computing resources used for the calculation of impulse responses and SEI are given. Simulation results are presented and discussed in Section 3. Finally, Section 4 provides a conclusion to the research findings and discusses future work.

2. Methods

Impulse responses and corresponding values of SEI were calculated for three shoebox-shaped rooms: a cube-shaped room, the room from Allen and Berkley's paper [1], [3], and its modified version. Their dimensions and volumes are given in Table 1. All three rooms were modelled with a frequency- and angle-independent reflection coefficient of all room boundaries set to 0.93. The image source method implementation given in [4] was used for the impulse responses generation, with the sampling frequency set to 22 050 Hz and the random distance parameter (used for the image source position randomisation) set to 0.

Table 1. Room dimensions and volume.

Room	Dimensions [m]	Volume [m^3]
Cube-shaped	[4, 4, 4]	64.00
Allen and Berkley's	[8, 12, 10]·343/800	75.66
Modified Allen and Berkley's	[8.3, 11.9, 10.2]·343/800	79.40

These three rooms serve to experimentally test the hypothesis formulated in [1]. The second room, therefore, in comparison with the cube-shaped room, ought to exhibit somewhat less pronounced echoes, while, in the third room, the sweeping echoes are expected to be weakest.

The measure of SEI utilised in this study is a modified, more robust version of the measure first proposed in [5]. In this modified version, a part of the early reflections region (specifically, the [50, 80] ms region after the arrival of the direct component) of the impulse response calculated with ISM up to the 10th reflection order, is first selected. Following that, the spectrum of this part of impulse response is calculated, and smoothed using a Gaussian filter. Finally, SEI is defined as the prominence of the strongest peak present in that spectral slice in the region above the first room mode (calculated using the formula for the rigid walls version given in [8]). The early reflections part is selected since the results presented in [5] indicated that the echoes remain slightly audible even when appearing only during that part of the impulse response. Unlike this version, in the original version of the SEI measure (used in [5]), the spectrum from which SEI was determined was based on a longer portion of the impulse response, but obtained with a substantially lower sampling frequency to achieve low compute times.

For each room, SEI was calculated for a set of eleven 2D grids of constant receiver height, with receiver positions discretized using a 2 cm step in both the x-axis and y-axis directions that spanned the horizontal plane. Receiver positions at room boundaries were not evaluated. These 2D grids were set at receiver heights in the [1.6, 1.8] m range in the z-axis direction, also with a 2 cm discretization step, in order to represent the common range of ground-up heights of an adult person's ear canal.

For each of the 2D grid receiver positions, a fixed set of 100 source positions was considered. Room coordinates-wise, the set of source positions was consistent across the three rooms and eleven receiver heights. It was comprised of 100 positions randomly chosen beforehand within an imaginary box with dimensions of [2, 2, 0.5] m, with a discretisation step of 1 cm set in all three dimensions. In each of the three rooms, the imaginary box with 100 source positions was placed 1 m from the room origin in both the x-axis and y-axis directions, and offset 1.5 m in the z-axis direction. In this manner,

a set of 100 source positions with the expected ground-up heights of the loudspeaker driver or an adult's mouth in the [1.5, 2.0] m range were considered.
The impulse responses generation and SEI calculation were performed on the HPC resources of the University of Rijeka. Each of the 50 utilized CPU cores, in a single simulation run, was assigned a 2D grid of receiver positions of predetermined height and one source position. To achieve this, Matlab code was exported to C language and compiled under the Ubuntu OS. SEI measure was calculated for a total of 3300 2D grids, providing 140 607 500 individual values. The average time for one impulse response generation plus SEI calculation, per core, was about 30 ms.

3. Results

The values of SEI obtained by simulations are presented in Figure 1, sorted in descending order. Overall, the data in panel a) support the hypothesis that stronger sweeping echoes occur when room dimensions are more regular, i.e. related by integer factors.

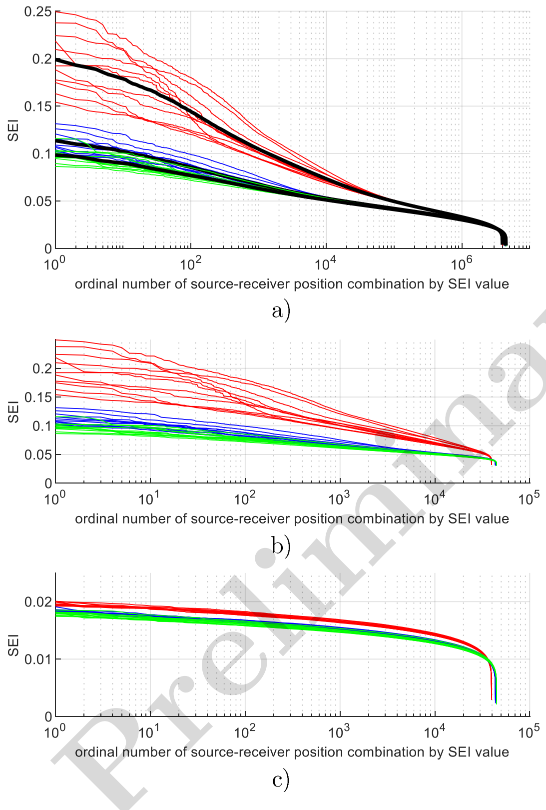

Figure 1. Values of SEI sorted in descending order: a) SEI values for all the source-receiver combinations, b) maximal SEI values for each receiver position on the 2D grid across the 100 source positions, c) minimal SEI values for each receiver position on the 2D grid across the 100 source positions. SEI values are presented separately for each 2D grid height. Black lines denote arithmetic means across the 11 heights. Rooms: cube-shaped room (red), Allen and Berkley's room (blue), and modified Allen and Berkley's room (green).

The worst-case scenario for each receiver position on a 2D grid is presented in panel b), for which, in total, 47 % of values are higher than 0.05. On the other hand, for the best-case scenario, where the lowest SEI value across the 100 source positions is selected for each receiver position, 83 % of values in total are within the [0.01, 0.02] range.
To quantify the highest SEI values present on a 2D grid, the 97.5$^{\text{th}}$ percentiles of the SEI values on individual grids, for a fixed source position, were calculated and are presented in Figures 2 and 3. In the former figure, it can be observed that the dispersion of the 97.5$^{\text{th}}$ percentiles across the 11 grid heights for a given source position is substantially larger for the cube-shaped room than for the remaining two rooms. In addition, for the first room, the arithmetic means of the 97.5$^{\text{th}}$ percentiles are about 0.05, whereas for the other two rooms the means are slightly above or around 0.04. Figure 3 shows a complementary view of the same data, indicating that the arithmetic mean of the 97.5$^{\text{th}}$ percentiles, calculated across the 100 source positions, changes only slightly across grid heights, regardless of room dimensions.

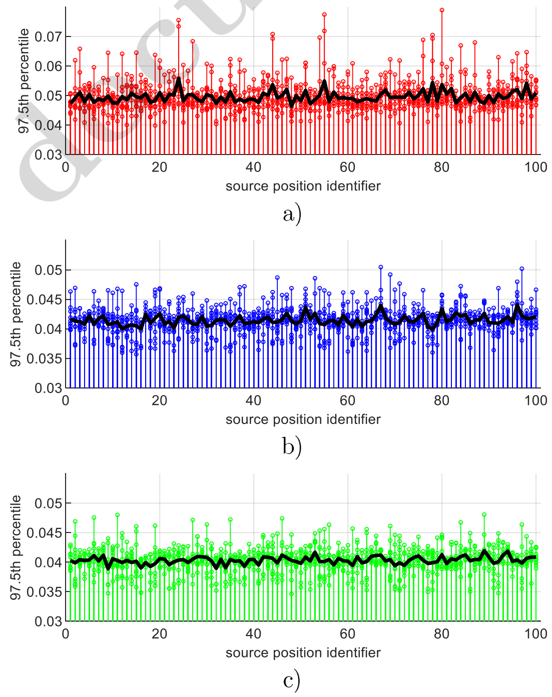

Figure 2. 97.5$^{\text{th}}$ percentile of SEI values on individual 2D grids for a fixed source position - grouped by individual source positions. Black lines denote arithmetic means. Rooms: a) cube-shaped room (red), b) Allen and Berkley's room (blue), and c) modified Allen and Berkley's room (green).

SEI patterns for a 2D grid set at 1.68 m height in the cube-shaped room that occur for three source positions are shown in Figure 4. Even though the panels present distinct patterns of sweeping echo intensities, they share similarly shaped circular patches of high intensity. Furthermore, the panel a) in Figure 5 and the panels in Figure 6 present SEI patterns for the worst-case scenario, i.e. the maximal value obtained for each receiver position, indicating that the higher SEI values are generally scattered across the grid.

With randomisation of the positions of image sources, regardless of the receiver position, the intensity of sweeping echoes becomes reduced – this is illustrated in the panel b) in Figure 5 for which the G-RIM method [9] was used. G-RIM is a modification of RIM, introducing randomisation of the image sources' positions within the volume enclosed by a sphere. The values of the azimuth and polar angles were sampled based on the procedure presented in [8], therein applied to ray-tracing for the creation of a spherically uniform radiation pattern.

To examine the audibility of the echoes present in the three rooms, an informal, non-blind listening test was performed by the study author (with normal hearing and ten years of formal music education). Short clean, i.e. "dry", speech utterances were convolved with a few hundreds of impulse responses with SEI values in the [0.015, 0.1] range, and listened via Sennheiser HD 550 headphones. For SEI values down to about 0.05, the sweeping echoes could clearly be perceived.

Based on this audibility threshold, in Table 2 the percentages of impulse responses with medium-to-weak ($0.05 \leq$ SEI ≤ 0.1) and strong (SEI > 0.1) sweeping echoes present in each room are summarized. Percentages indicate that even though the strong sweeping echoes were detected 745 times more frequently in the cube-shaped room than in the modified Allen and Berkley's room, the detection-frequency ratio reduced to 8.1:1 for the medium-to-weak echoes.

Finally, Figure 7 illustrates the two extremes – a case with a high SEI value (0.206), and a case with a low SEI value (0.037), both for impulse responses obtained for the cube-shaped room. In the first case, sweeping echoes are clearly visible – in the time-frequency domain as straight lines of high intensity (panel a)), in the impulse response as an impulse train of gradually diminishing time delays, and in the energy decay curve (EDC) as a staircase (panel b)). On the other hand, for the second case, no such regularities can be observed (panels c) and d)).

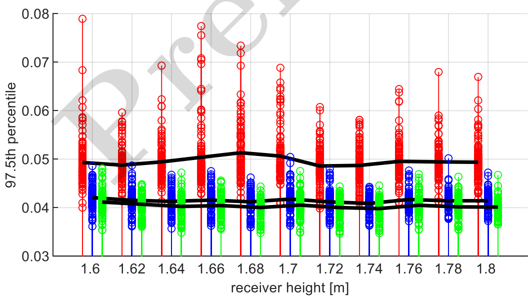

Figure 3. 97.5th percentile of SEI values on individual 2D grids for a fixed source position - grouped by individual grid heights. Black lines denote arithmetic means. Rooms: cube-shaped room (red), Allen and Berkley's room (blue), and modified Allen and Berkley's room (green).

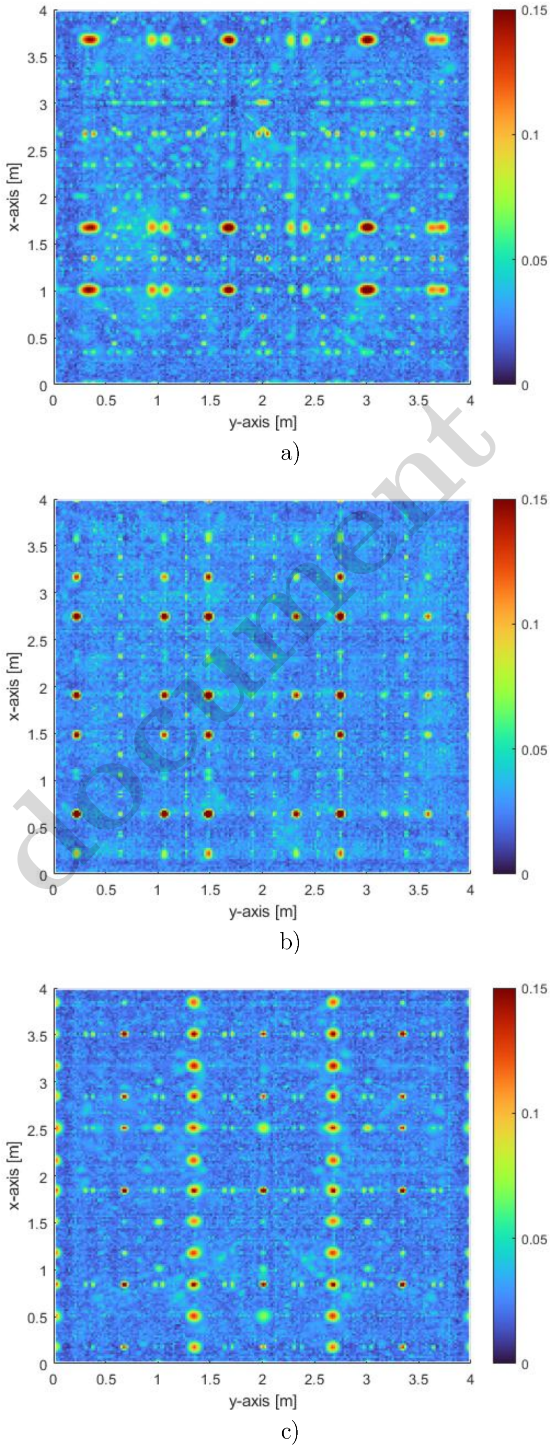

Figure 4. Sweeping echo intensity patterns for a 2D grid set at 1.68 m in the cube-shaped room: a) source position: [2.99, 1.71, 1.67] m, b) source position: [2.51, 1.26, 1.89] m, c) source position: [1.50, 1.98, 1.66] m. At specific receiver positions, the values surpass the upper colorbar limit.

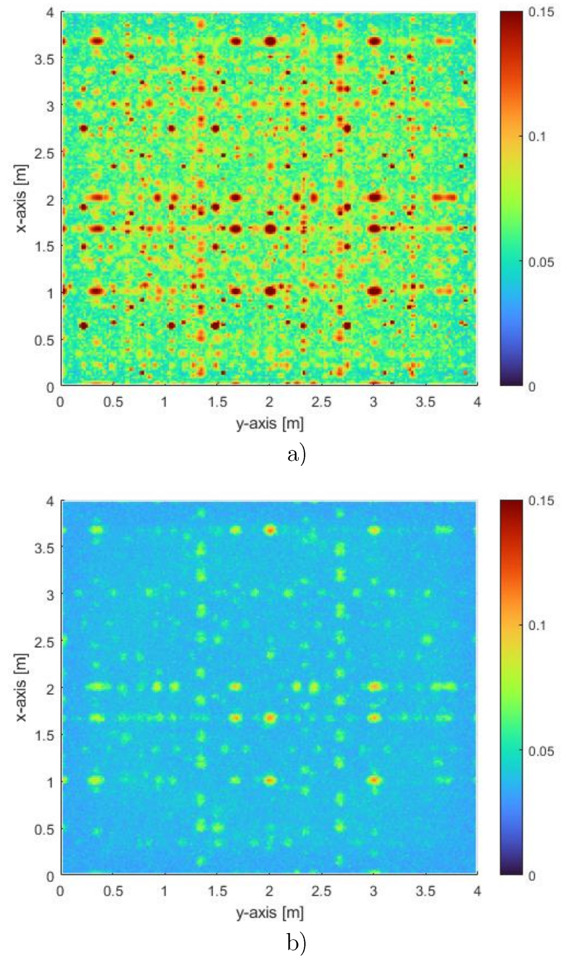

Figure 5. Sweeping echo intensity patterns for a 2D grid set at 1.68 m in the cube-shaped room - maximal SEI values across the 100 source positions: a) no randomisation of the positions of image sources, b) G-RIM randomization using a sphere radius of 7.4 cm. At specific receiver positions in panel a), the values surpass the upper colorbar limit.

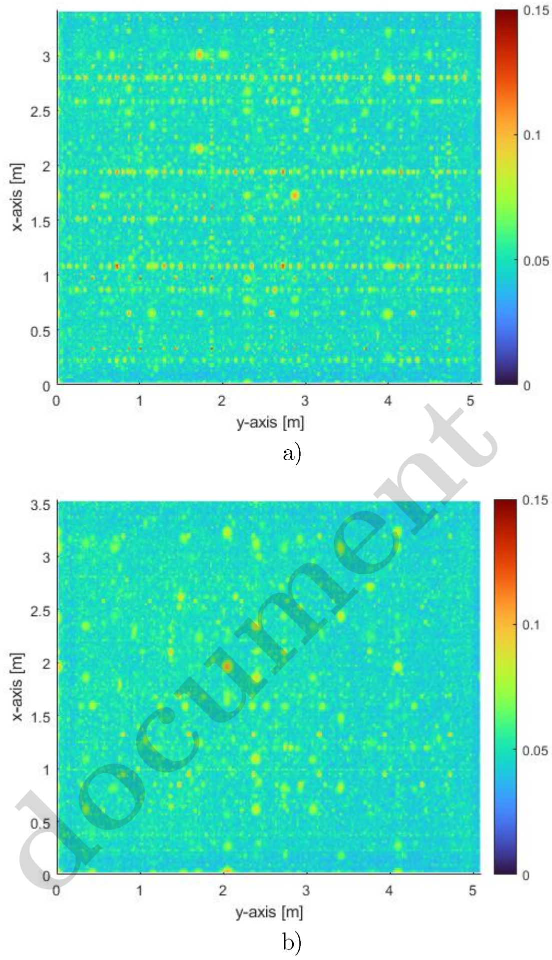

Figure 6. Sweeping echo intensity patterns for a predetermined 2D grid height - maximal SEI values across the 100 source positions: a) Allen and Berkley's room with grid height set at 1.72 m, b) modified Allen and Berkley's room with grid height set at 1.60 m.

Table 2. Percentage of SEI values within two ranges. Subscript 'Max' denotes calculations based on the 2D grids for which maximal SEI values across the 100 source positions have been used, whereas 'All' indicates all simulation data have been used.

Room/ SEI range	Cube-shaped	Allen and Berkley's	Modified Allen and Berkley's
$[0.05, 0.1]_{All}$	2.35 %	0.41 %	0.29 %
$(0.1, +\infty)_{All}$	$3.77 \cdot 10^{-2}$ %	$4.64 \cdot 10^{-4}$ %	$5.08 \cdot 10^{-5}$ %
$[0.05, 0.1]_{Max}$	84.14 %	32.44 %	24.39 %
$(0.1, +\infty)_{Max}$	3.60 %	$4.64 \cdot 10^{-2}$ %	$5.08 \cdot 10^{-3}$ %

4. Conclusion

The results of this computational study indicate that sweeping echoes are present and, based on an informal listening test, audible at various source-receiver positions not only in shoebox-shaped rooms with regular parameters. Moreover, simulation results revealed that the echoes of moderate-to-weak intensity are only about eight times more frequent in a cube-shaped room than in a room with irregular room dimensions. It was also demonstrated that the echoes can be substantially reduced or completely eliminated, regardless of the source or receiver position, by randomisation of the positions of image sources. A relevant direction for future research includes the evaluation of sweeping echo intensity patterns in shoebox-shaped rooms of smaller volume and of lower ceiling height, commonly found in dwellings.

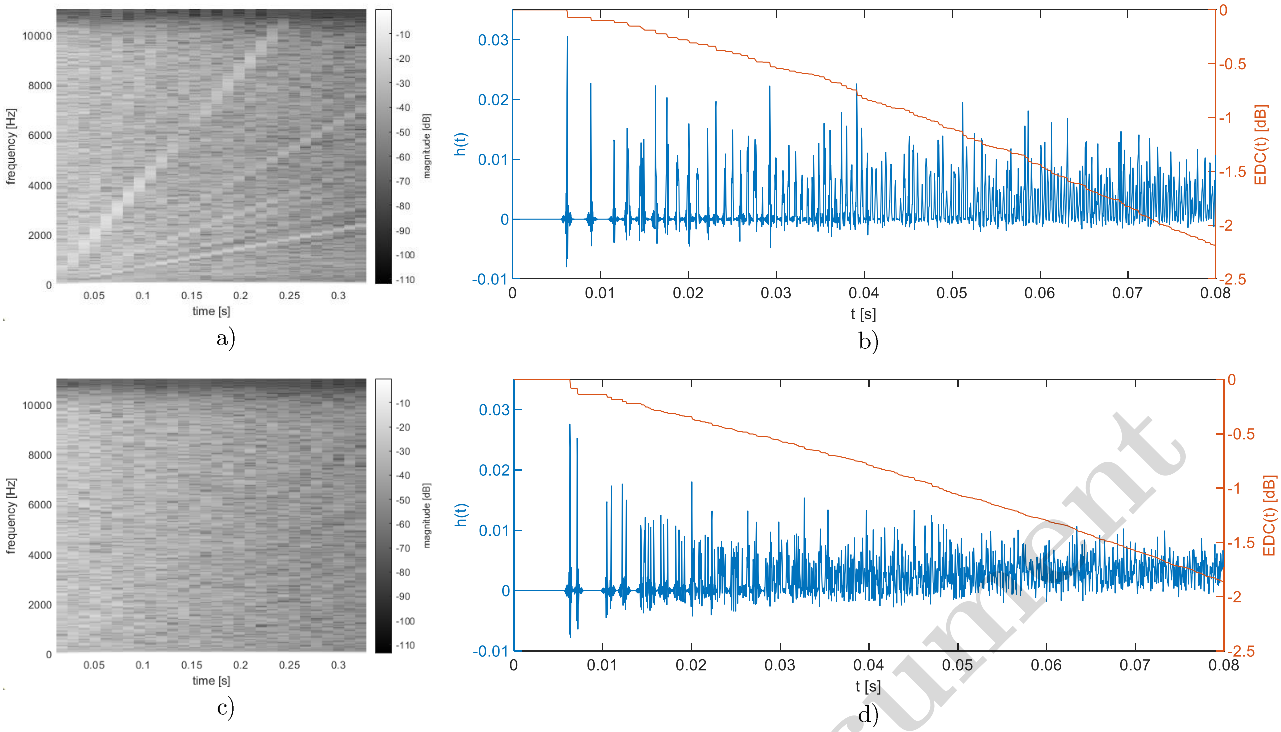

Figure 7. Spectrograms and time domain representations of the first 80 ms of impulse responses and of their corresponding energy decay curves (EDCs) for two source-receiver positions. Impulse responses were obtained for the cube-shaped room. Panels a) and b) – source position: [1.66, 2.50, 1.68], receiver position: [2.32, 0.48, 1.66], SEI: 0.206. Panels c) and d) – source position: [1.99, 2.08, 1.93], receiver position: [3.84, 0.98, 1.66], SEI: 0.037.

5. Acknowledgments

The author gratefully acknowledges the Center for Advanced Computing and Modelling of the University of Rijeka for providing high-performance computing resources. The author is especially thankful to Dr. Stefan Ivić and Dr. Ante Sikirica for their kind support and technical guidance.

References

[1] E. de Sena, N. Antonello, M. Moonen, and T. van Waterschoot: "On the modeling of rectangular geometries in room acoustic simulations," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 4, pp. 774-786, 2015.

[2] O. Das, and E. De Sena, "The Complex Image Method for Simulating Wave Scattering in Room Acoustics," in *Proc. of Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, (Bologna, Italy) , pp. 1-7, 2023.

[3] J. B. Allen, and D. A. Berkley: "Image method for efficiently simulating small-room acoustics," J. Acoust. Soc. Amer., vol 65, no. 4, pp. 943–950, 1979.

[4] https://github.com/enzodesena/rim Last accessed 2026-03-28.

[5] A. Andrijašević, and M. Horvat, "Sweeping echoes management in image-method based open-source software," in *Proc. of 11th Convention of European Acoustics Association – Forum Acusticum-Euronoise*, (Malaga, Spain), pp. 4193-4199, 2025.

[6] https://github.com/Building-acoustics-TU-Eindhoven/OSS_RoomAcoustics/blob/main/Open_Software_Room_Acoustics_overview.md Last accessed 2026-03-28.

[7] M. C. J. Hornikx, H. Wang, I. Fichera, L. Paganini, A. Nolte, and A. Serebrenik, "Exploring the Current Landscape of Open Research Software in Room Acoustics," in *Proc. of 53rd International Congress and Exposition on Noise Control Engineering (Inter-noise)*, (Nantes, France), pp. 5875-5884, 2024.

[8] M. Vorländer: *Auralization*. Springer Cham, 2020.

[9] A. Andrijašević, "Gaussian distribution-based randomised image method," in *Proc. of 53rd International Congress and Exposition on Noise Control Engineering (Inter-noise)*, (Nantes, France), pp. 8478-8488, 2024.