An Analysis of the Computational Efficiency Gains from the Incorporation of Importance Sampling in Determining Optimal Sound Source Location

Joshua Azimullah

Supervisor(s): Dr. Jorge Abraham Martínez Castañeda, Prof. Dr. Elmar Eisemann

[1]EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
January 29, 2023

Name of the student: Joshua Azimullah
Final project course: CSE3000 Research Project
Committee: Dr. Jorge Abraham Martínez Castañeda, Prof. Dr. Elmar Eisemann, Prof. Dr. Marcus Specht

Abstract

A novel approach for determining the optimal location of a sound source within an acoustic environment is proposed. This approach involves the application of Importance Sampling to improve the efficiency of the existing method of acoustic ray-tracing for finding the frequency response at various listening locations. The results of this study do not conclusively demonstrate the superiority of the proposed method over existing techniques, and further research is necessary to determine its viability.

1 Introduction

This thesis presents a more efficient methodology for identifying the optimal placement of a sound source within a given acoustic environment. The objective of this study is to determine the location of the sound source that results in the highest acoustic quality, as perceived by listeners, at predetermined receiver positions. Since the determination of what constitutes "acoustic quality" is inherently subjective, a more objective measure, namely spectral flatness [Dub04], is employed in this study. The measurement of spectral flatness assesses the level of flatness in the frequency response. It can be argued that a higher level of spectral flatness results in a more favorable auditory experience, one that more closely aligns with the artist's intent. In this research, the highest average spectral flatness is utilized as the primary metric for determining the optimal acoustic quality of a given space.

In order to determine the frequency response at receiver positions, a Room Impulse Response (RIR) must be generated from a three-dimensional model of the space under analysis. RIRs serve as a representation of the reverberant acoustic characteristics of a given space. They can be utilized to replicate the reverberant effects in a simulated environment or evaluate the acoustic qualities of a given space. The process of creating RIRs from three-dimensional models of an acoustic environment is referred to as auralization [SS15; kle93; Sch11].

This thesis presents a methodology for identifying the optimal placement of a sound source within a given acoustic environment through the use of a more efficient variation of acoustic ray-tracing, itself an auralization technique. In acoustic ray-tracing, RIRs are generated by tracing acoustic rays through a 3D model of a given space and simulating the sound reflections and transmissions at each surface. The subsequent detection of these rays is accomplished through the utilization of a receiver sphere, wherein the energies of the rays traversing it are quantified and recorded. The number of rays necessary to adequately penetrate the receiver sphere in complex models, such as those with a high inner volume and a large number of triangles, rapidly increases. However, casting a large quantity of rays can lead to significant computational expense [VS08].

This thesis proposes the integration of Importance Sampling (IS) within acoustic ray-tracing. IS is a statistical technique that focuses on sampling the regions of high probability density in order to improve the accuracy and efficiency of numerical integration and estimation. IS techniques are applied to selectively cast rays in the most useful directions, such that most rays go through the receiver. The use of IS could enable the reduction of computational requirements while maintaining an acceptable level of accuracy in the RIRs in order to determine the optimal sound source location.

The methodology employed in this study involves the implementation of an acoustic ray-tracing simulation, similar to the one described in [Tho17]. The simulation was then adapted to incorporate the use of IS for determining the direction of initial rays cast from the source. The efficiency of the IS-augmented simulation is evaluated by comparing its computational speed with that of the non-IS simulation, as well as assessing the similarity of their spectral flatness.

In the background information section, a brief overview of previously employed auralization methods is discussed. Subsequently, a preliminary overview of the process of acoustic ray-tracing, its conversion to a RIR and spectral flatness determination is provided. In the proposed methodology section, the implementation of IS techniques are described in detail. The criteria for determining the optimal paths for the rays is discussed, as well as the methods used to construct the sampling distribution. In the following section, an in-depth examination of the simulation process, implementation details, and the parameters used in the study is discussed. The section on responsible research addresses any ethical considerations and adherence to the principles of F.A.I.R. research. In the results section, the performance differences and flatness similarities between the IS-augmented and non-IS simulations is presented and analyzed. The study specifically investigates the computational efficiency and the accuracy of the spectral flatnesses generated under different starting parameters. The discussion section includes an examination of the results and proposes avenues for future research. Finally, the conclusion summarizes the key findings of the study and provide insights into the potential applications and implications of the

research. The conclusion also discusses the limitations of the study and suggest areas for future research.

2 Background information

Previous research has employed a range of auralization techniques [Tho17; Cas13; SS15; SSMS13; TWH+22; HN06; KSS68; HH84]. The auralization techniques can be divided in two categories: Geometric acoustics and wave-based modeling [Cas13; SS15; SSMS13; Tho17]. Wave-based modeling utilizes the wave properties of sound to create an auditory representation of a virtual environment [Cas13]. Wave-based methods are not suitable for higher frequencies since the wave equations used become complex, making them difficult to compute [MI86; SSMS13]. Additionally, the level of detail obtained exceeds that which is necessary for making accurate assessments of the room's acoustical characteristics [Cas13; Kut14]. Instead, geometric methods, similar to those used in geometric optics, can be utilized. These can be used where the dimensions of the space are significantly larger than the wavelength of the sound waves [Kut14] . This approach is simpler and more computationally efficient [SS15; Cas13]. This thesis employs the technique of acoustic ray tracing, as outlined in various literature sources [Sch11; Kut14; SS15; KSS68]. This method is particularly beneficial for the analysis of complex geometries, as demonstrated in [Cas13; Bor84]. Additionally, the ability of this technique to model diffusivity, as highlighted in [SS15], further supports its utility in this context.

3 Acoustic ray-tracing Simulation and Spectral Flatness Analysis

In order to identify the optimal source location for an enhanced auditory experience, predetermined source and receiver locations must be established. The spectral flatness is then calculated for each distinct combination of source and receiver locations. The source location yielding the highest average spectral flatness across all receiver locations is determined to be the optimal location.

To calculate the spectral flatness of a given source-receiver combination, an acoustic ray-tracing simulation is conducted. The following sections of this chapter provide a detailed examination of the specific methodology employed in the acoustic ray-tracing simulation utilized in this thesis, as well as the procedures for determining the spectral flatness following the simulation.

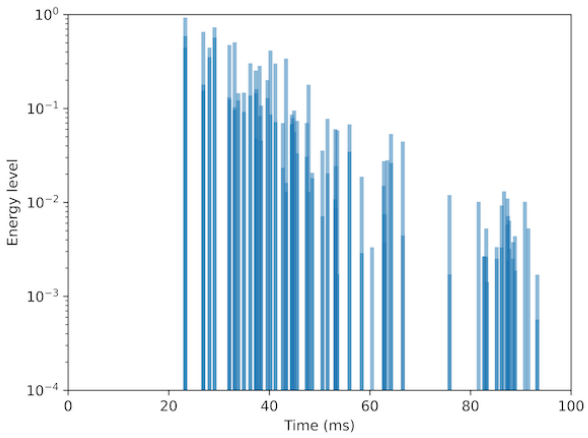


Figure 1: Histogram showing the amount of energy received over time

3.1 Acoustic ray-tracing

Acoustic ray-tracing models sound waves as sound rays [Kut14; Sch11; Cas13]. They are emitted from a point representing the sound source location [Sch11; Cas13; SS15]. The initial directions of these rays are generated using the method outlined in Section 4.

Each ray is assigned an initial energy vector $\vec{\mathrm{E}}$, wherein the magnitude of energy of several frequency bands is presented [Tho17]. Upon initiation, the energy of each frequency band is initialized to a value of one, in order to simulate a spherical uniform audio source.

Upon impact with a surface, the energy vector $\vec{\mathrm{E}}$ of a ray is attenuated by $\vec{\mathrm{E}}_{\mathrm{outgoing}} = \vec{\mathrm{E}}_{\mathrm{incoming}} \cdot (1 - \vec{\mathrm{S}}) \cdot (1 - \vec{\alpha})$, similar to the methods described in [Tho17; SSMS13; Kut14; Sch11; SS15; VM00]. In this equation, $\vec{\mathrm{E}}_{\mathrm{outgoing}}$ represents the energy vector of the outgoing ray, $\vec{\mathrm{E}}_{\mathrm{incoming}}$ represents the energy vector of the incoming ray, and the vectors $\vec{\mathrm{S}}$ and $\vec{\alpha}$ represent the scattering and absorption coefficients, respectively. The absorption coefficients are used to quantify the amount of energy absorbed by a given surface. The scattering coefficients, as discussed in Section 3.2, are used to quantify the amount of energy that is not directly specularly reflected, but rather scattered uniformly throughout space.

Following the impact with a surface, a new reflected ray is generated with the attenuated energy vector. This process is repeated until a predetermined maximum number of impacts is reached.

3.2 Histograms and Energy detection

Histograms offer the capability to document the energy detection over a period of time in a sequence of energy bins,

providing a representation of the energy received [Sch11; KSS68; Kut14]. An example histogram is shown in Figure 1. For each frequency band, a separate histogram is created to depict the energy received by that specific frequency band, providing an analysis of the energy received at different frequencies.

The energy recorded in the histogram includes both specular and scattering energies, which are subsequently adjusted by $1/r^2$, where r represents the total distance traversed by the rays from the source to reach the receiver. This approach is necessary to account for the inverse square law, which states that the intensity of the wave decreases in proportion to the square of the distance from the source [Sch11].

Specular energies

To accurately record the specular energies, all rays that have passed through the receiver sphere are identified. The energies of these rays, after being adjusted for the inverse square law, is added to the appropriate frequency band's histograms at the time $t = \frac{d}{c}$, where t represents time, d represents the total distance from the source, and c represents the speed of sound.

Diffuse energies

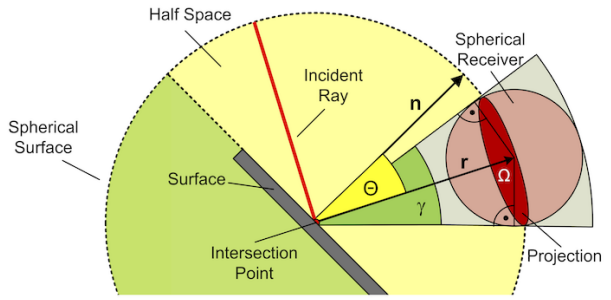


Figure 2: The following image depicts the amount of diffuse energy recorded onto the histogram, where $\mathbf{n}$ represents the surface's normal, $\mathbf{r}$ the distance from the intersection point to the receiver, and γ represents the angle of projection Ω from the receiver sphere on the hemisphere. Image source: [Sch11].

The method of diffuse rain, as discussed in reference [Sch11], is utilized to document the scattering energies emitted by rays. This methodology assumes that the scattering energies can be modeled as the propagation of a lambertian distribution of energy emitted from the impact points from all rays. In other words, it postulates that each intersection point of a ray emits a hemisphere of diffusional energy over the entire space.

The diffuse energy recorded in the histogram is based on both the proportion of incoming ray energy that is dispersed as diffuse energy, as determined by the scattering coefficient, and the ratio of the area of the receiver sphere that is projected onto the total diffuse hemisphere, as depicted in Figure 2.

The diffuse energy is only added to the histogram if a ray from the intersection point has a direct line of sight to the receiver. The energy added to the histogram as a result of diffuse scattering is determined by the equation outlined in [Sch11]. However, the values of Θ and γ, as depicted in Figure 2, are implicitly incorporated in the following equation: $\vec{\mathrm{E}}_{\mathrm{diffuse}} = \vec{\mathrm{E}}_{\mathrm{incoming}} \cdot (1 - \frac{r_{\mathrm{receiver}}}{r_{\mathrm{hemisphere}}}) \cdot 2(d_{\mathrm{diffuseRay}} \cdot n_{\mathrm{surface}})$, where vector $\vec{\mathrm{E}}_{\mathrm{diffuse}}$ represents the diffuse energy, vector $\vec{\mathrm{E}}_{\mathrm{incoming}}$ represents the incoming energy, r_{receiver} represents the radius of the receiver sphere, and $r_{\mathrm{hemisphere}}$ represents the distance to the receiver. The dot product, $d_{\mathrm{diffuseRay}} \cdot n_{\mathrm{surface}}$, represents the cosine angle between the diffuse ray and the surface normal. Finally, the energy is added to the histogram at the time $t = \frac{d_{\mathrm{total}} + r_{\mathrm{hemisphere}}}{c}$, where d_{total} represents the total distance traveled by the ray from the source to the intersection point, and c represents the speed of sound.

3.3 Auralization

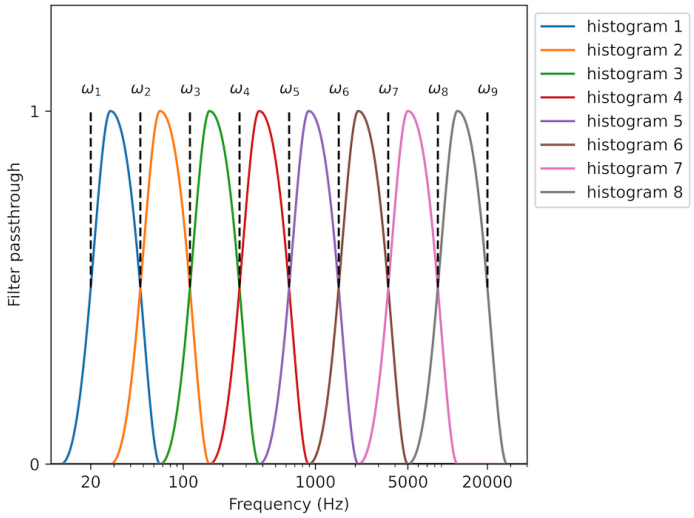


Figure 3: Frequency bands used to filter the histogram bands, where ω_i represent the frequency band edges

The histograms generated in this thesis have a sampling rate of 44100 Hz. These histograms can be readily transformed into an audio waveform file through a process of normalization and alternate weighting. Each bin in the histogram corresponds to a sample in the resultant audio file. Normalization is applied to the histograms prior to conversion, as it ensures that the overall volume level of the audio file is not

affected by the number of rays cast during the simulation. This is done, since the focus in this thesis is on the variations in volume across the frequency spectrum. The normalized histograms are then transformed into samples by applying an alternate weighting of 1 and -1 at each time-step, and then discretizing these values onto a 16-bit signed integer.

Prior to combining the audio samples generated from the histograms, each sample undergoes filtering utilizing both high-pass and low-pass filters. The high-pass filter eliminates all frequencies below a specified cutoff, while the low-pass filter eliminates all frequencies above a specified cutoff. These filters have a defined edge frequency, which marks the crossover point between two frequency bands. The edge frequencies for the histograms in this thesis are determined using the following equation:

$$\omega_{\text{edge}_i} = \omega_{\text{lowest}} \left(\frac{\omega_{\text{highest}}}{\omega_{\text{lowest}}} \right)^{\frac{i}{N_{bands}}}, i = 1, \ldots, 9$$

In this equation, ω_{lowest} and ω_{highest} represent the lowest and highest frequencies of 20 Hz and 20 kHz, respectively, and N_{bands} represents the number of frequency bands used, which is in this thesis defined as 8 bands, similarly to [Tho17] . The i-th histogram utilizes the frequency edge ω_{edge_i} for its high-pass filter and frequency edge $\omega_{\text{edge}_{i+1}}$ for its low-pass filter. The defined frequency bands can be observed in Figure 3.

The filters selected for this thesis are described in [Ant10]. The filters which are chosen for this research from that reference, are designed to ensure an overlap that maintains a constant sum, resulting in a natural sounding overlap [Tho17].

Finally, the RIR, is constructed by summing all of the previously filtered samples. This resultant sample can then be further analyzed.

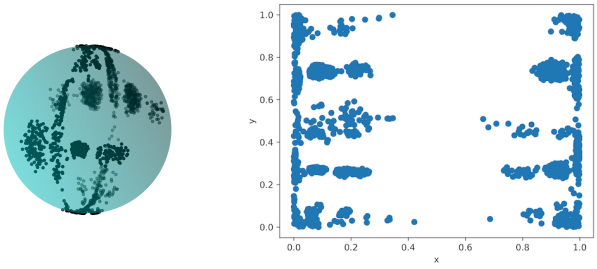
3.4 Spectral flatness analysis

Subsequently, the spectral flatness is computed from the produced RIR. The spectral flatness serves as an indicator of the uniformity of the frequency response. The measure of spectral flatness is characterized by the ratio of the geometric mean to the arithmetic mean of the power spectrum, which is obtained by squaring the absolute values of the Fast Fourier Transform of the signal [Dub04; Coo16]. The equation for spectral flatness is represented as:

$$\text{SF} = \frac{\exp(\sum_{i=1}^{n-1} x_i)}{\frac{1}{n} \sum_{i=1}^{n-1} x_i}$$

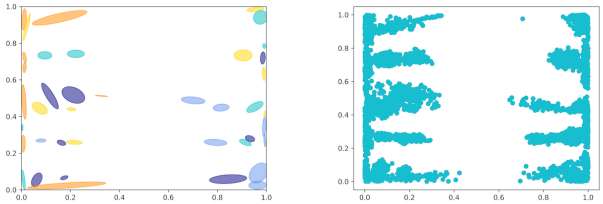
Where SF is the spectral flatness, x_i are the values of the power spectrum and n is the number of values in the power spectrum.

4 Proposed methodology for generating initial ray directions



(a) (b)

Figure 4: Initial ray directions depicted as points on a 3D unit sphere and their projected points in 2D



(a) (b)

Figure 5: Distributions generated from previous ray directions (a), Newly sampled points from gaussian distributions (b)

In this chapter, a novel method for acoustic ray-tracing that utilizes Importance Sampling (IS) through the implementation of gaussian mixture models (GMM) [RLMS+18] is presented. Specifically, a new probabilistic method for generating initial unit rays is proposed.

A Gaussian Mixture Model (GMM) is a model that is based on the assumption that the input data is generated from a mixture of several Gaussian distributions. The data used for training the GMM distributions are initial ray directions that pass

through the receiver when traced. These distributions generated are then leveraged to generate new, more probable directions for the ray-tracing process.

Rays are initially cast uniformly over a sphere in accordance with traditional methods. However, this is done with a reduced number of rays to achieve the proposed performance gains. The directions of rays that subsequently pass through the receiver, after the ray-tracing process, are then marked and projected onto a 2-dimensional plane, as depicted in Figure 4. The projection utilizes the equidistant projection method outlined in reference [Wei95a]. A projection is employed due to the computational difficulties associated with generating random distributions in 3-dimensional space, as well as the fact that a significant portion of the probability density function would otherwise be unused, given that all initial ray directions are unit vectors.

The projected 2-dimensional points are then used as input for the GMM. This results in the generation of multiple distributions, each with corresponding means and covariances, as illustrated in Figure 5a. In this thesis, 25 of such components are generated. The GMM is then used to sample new 2-dimensional points, which are more likely to hit the receiver, as illustrated in Figure 5b. These points are then converted back into 3-dimensional initial vectors using the method seen in [Wei95a; Wei95b]. Subsequently, in order to mitigate the potential for over-saturation of the receiver from directions with high ray density, while also ensuring that directions with low ray density are not overlooked, the energy cast by each ray is adjusted through the application of the inverse probability of generating the corresponding 2-dimensional point. After this adjustment, the acoustic ray-tracing is performed over the newly generated directions.

5 Experimental setup and Implementation

5.1 Experimental setup

Models

In the present study, simulations are performed in a small room, a theatre, and an industrial warehouse. The first two environments were chosen as they are representative of typical scenarios where the determination of an optimal source location is desired. The industrial warehouse was selected as it represents an asymmetric structure, which is of particular relevance to the current thesis. The models utilized in this study can be found in the appendix section.

Simulations

In Section 7, an assessment of the effectiveness of the proposed method is conducted by comparing a simulation with a relatively high number of rays cast to several simulations with lower ray counts, either employing IS or not. The comparison is performed using simulations consider specular energies, diffuse energies, and both. Subsequently, a comparison between the computational performance in terms of the number of rays cast. The number of rays used for the high-number simulation is 100,000,000 and is called the baseline simulation.

Consistent absorption and scattering coefficients were employed across all tests. Further investigations are required to evaluate the impact of varying coefficients on the results.

Error metric

The spectral flatness of each generated RIR is calculated and smoothed using a Hann window of 1000 in order to assess the overall frequency spectrum, as opposed to any potential noise. The absolute difference between the flatness of the high ray count simulation and the flatness of each corresponding simulation is then computed, serving as an error metric for the simulation.

5.2 Implementation

The implementation of the system is divided into two phases: ray-tracing and auralization. The ray-tracing component, which involves reflections and energy recording, was implemented in C++ for performance optimization. Each ray was evaluated in parallel and the maximum number of reflections was set to 5. Following the creation of a histogram, the analysis, including the generation of the power spectrum and the calculation of the error metric, was performed using Python.

The Gaussian Mixture Model was implemented using the scikit-learn library [PVG+11]. The C++ implementation was integrated with the Python direction generation module to apply IS.

All simulations were conducted on a MacBook Pro equipped with an M1 Pro chip and 32GB of RAM.

6 Responsible Research

The ethical considerations of this study are limited, as it primarily involves physical data from audio and simulations, rather than personal or sensitive information. However, when applying the results to real-world scenarios, such as predicting speaker locations in private homes, it is important to consider privacy and data protection. Additionally, the use of resources for such purposes should be evaluated in light of en-

vironmental impact and alternative areas where the resources may be better utilized. Finally, it is recommended that users approach the results of this study with caution and critically evaluate the validity and accuracy of the models. The code used in this study is available online to promote transparency and reproducibility.

7 Results

Rays	Specular		Diffuse		All	
	No IS	IS	No IS	IS	No IS	IS
1,000,000	5.4	19.2	3.8	148.8	1.8	64.7
100,000	12.8	0.8	64.0	149.8	51.6	39.0
50,000	25.2	5.8	78.6	156.2	64.1	42.9
10,000	63.4	45.1	87.0	58.1	96.8	20.9

(a) Room model comparison

Rays	Specular		Diffuse		All	
	No IS	IS	No IS	IS	No IS	IS
1,000,000	4.0	2.8	12.3	177.0	99.3	218.0
100,000	40.2	178.8	83.7	290.3	243.8	315.0
50,000	47.6	47.8	79.2	276.1	250.5	218.2
10,000	47.7	109.4	62.0	92.2	228.3	232.6

(b) Theater model comparison

Rays	Specular		Diffuse		All	
	No IS	IS	No IS	IS	No IS	IS
1,000,000	10.9	30.6	163.0	138.4	24.4	16.3
100,000	76.1	170.9	195.6	189.2	169.1	76.2
50,000	171.5	41.9	152.5	199.0	243.0	16.9
10,000	317.7	247.5	174.2	194.8	303.3	136.7

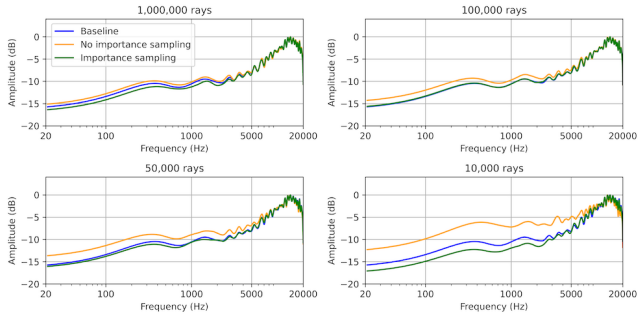
(c) Warehouse model comparison

Table 1: Comparing the absolute difference in spectral flatness against a simulation with high ray count for both Importance Sampling (IS) and No Importance Sampling (No IS) (x1000)

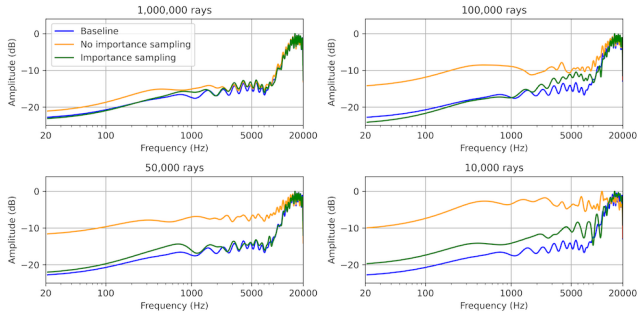
In Table 1 the errors are presented in relation to the high-ray-count simulation of 100,000,000 rays. It is observed that, in the majority of instances where IS is not applied, the error decreases as the number of rays increases, as expected.

It is noteworthy that the results of the IS simulation indicate that, for some models, there is an optimal range where the simulation results are significantly closer to the baseline than those obtained from non-IS. However, this is not always the simulation that utilizes the highest number of rays. This is most evident in the specular only energy power spectrum of the room model, shown in Figure 6a, where the IS simulation outperforms the non-IS simulation by a substantial margin at 100,000 rays.

In general, it appears that all IS simulations are significantly deviant in terms of diffuse energy values. This highlights the need for further refinement in the simulation techniques and modeling approach in order to accurately predict



(a) Specular only energy comparison in the room model



(b) Specular and Diffuse energy comparison in the warehouse model

Figure 6: A comparison of the power spectra between a high-ray-count simulation without importance sampling, the baseline, and simulations with varying ray counts, both with and without importance sampling, recording both specular and diffusion energies

diffuse energy in these scenarios.

The examination of both specular and diffuse energies, as depicted in Figure 6b for the warehouse model, reveals that the IS simulation with 50,000 rays, generally follows the same patterns as the baseline. The deviation might be due to an improper adjustment of probabilities. It is worth noting that without any probability adjustment, the power spectra would be completely misaligned. Further investigation is necessary to determine the optimal probability adjustment strategy for IS.

The examination of Table 2 reveals that the use of IS results in a significantly slower performance in casting the same number of rays. Therefore, a substantial improvement in quality is necessary for its practical usage. Although there is potential for a significant increase in quality, currently it remains a slower method compared to non-IS.

8 Discussion and Future work

The proposed method may not yield expected results due to several factors. Firstly, it could be that the probability attenuation is incorrect and alternative balancing methods are necessary. Secondly, it may be that the distributions within

Rays	Specular		Diffuse		All	
	No IS	IS	No IS	IS	No IS	IS
1,000,000	999	25950	948	8931	1328	9421
100,000	337	1993	313	1934	467	2115
50,000	283	1481	311	1931	458	2080
10,000	244	1018	254	1093	388	1231
1,000	241	885	246	905	377	1036

(a) Room model performance (ms)

Rays	Specular		Diffuse		All	
	No IS	IS	No IS	IS	No IS	IS
1,000,000	14859	39417	31792	49981	32205	50503
100,000	1698	6291	3147	6458	3321	6645
50,000	953	2497	1741	3440	1898	3605
10,000	402	1461	570	1381	715	1520
1,000	265	970	302	930	430	1060

(b) Theater model performance (ms)

Rays	Specular		Diffuse		All	
	No IS	IS	No IS	IS	No IS	IS
1,000,000	30128	72952	66808	107781	67198	108241
100,000	3154	8513	7236	11303	7404	11475
50,000	1841	5026	3344	6085	3494	6242
10,000	551	2100	974	1844	1141	1985
1,000	286	901	327	1066	459	1198

(c) Warehouse model performance (ms)

Table 2: Duration of Simulation Completion in Milliseconds

the Gaussian Mixture Model (GMM) were too stringent, and a more relaxed and randomized approach with a differing number of components should be explored in future research. Thirdly, capturing diffuse rain energies may be inherently challenging as the diffuse rain requires random intersections points scattered throughout the model, which the IS method may miss. Further research is needed to determine if these limitations prevent the IS method from being applicable.

Additionally, the method of comparing to a high-ray simulation is not the optimal approach for evaluating performance. Future research could consider comparing the IS method to the real-world frequency response of a given room, using various coefficients and models for a more accurate assessment.

The determination of the optimal source location in this study was restricted by the computational requirements associated with utilizing such a high-ray simulation to examine a sufficient number of source-receiver combinations. The time required to accurately determine the optimal location was a constraint in this study. Nevertheless, a comprehensive analysis was performed with the resources available and important findings were still obtained.

9 Conclusion

In conclusion, it is currently unclear whether Importance Sampling (IS) is a suitable method for determining the optimal speaker location in a given environment. Further research is necessary to determine and proof the correct method of energy attenuation. At present, the limitations and performance degradation of IS in comparison to established acoustic ray-tracing methods result in inferior performance of IS for locating sound sources.

References

[Ant10] J. Antoni, "Orthogonal-like fractional-octave-band filters," *The Journal of the Acoustical Society of America*, vol. 127, no. 2, pp. 884–895, 2010, ISSN: 0001-4966. DOI: 10.1121/1.3273888. [Online]. Available: https://dx.doi.org/10.1121/1.3273888.

[Bor84] J. Borish, "Extension of the image model to arbitrary polyhedra," *The Journal of the Acoustical Society of America*, vol. 75, no. 6, pp. 1827–1836, 1984, ISSN: 0001-4966. DOI: 10.1121/1.390983. [Online]. Available: https://doi.org/10.1121/1.390983.

[Cas13] J. A. Castañeda, "Low-complexity computer simulation of multichannel room impulse responses," Doctoral Thesis, 2013. [Online]. Available: https://doi.org/10.4233/uuid:a3366a3f-1a76-4614-a4cf-fba7a0808ced.

[Coo16] J. D. Cook, *Spectral flatness*, Web Page, 2016. [Online]. Available: https://www.johndcook.com/blog/2016/05/03/spectral-flatness/.

[Dub04] S. Dubnov, "Generalization of spectral flatness measure for non-gaussian linear processes," *IEEE Signal Processing Letters*, vol. 11, no. 8, pp. 698–701, 2004, ISSN: 1558-2361. DOI: 10.1109/LSP.2004.831663.

[HH84] P. S. Heckbert and P. Hanrahan, "Beam tracing polygonal objects," *SIGGRAPH Comput. Graph.*, vol. 18, no. 3, 119–127, 1984, ISSN: 0097-8930. DOI: 10.1145/964965.808588. [Online]. Available: https://doi.org/10.1145/964965.808588.

[HN06] M. Hodgson and E. M. Nosal, "Experimental evaluation of radiosity for room sound-field prediction," *J Acoust Soc Am*, vol. 120, no. 2, pp. 808–19, 2006, ISSN: 0001-4966 (Print) 0001-4966. DOI: 10.1121/1.2216559.

[kle93] m. kleiner, "Introduction volume 41, number 11, on auralization," *journal of the audio engineering society*, vol. 41, p. 860, 1993.

[KSS68] A. Krokstad, S. Strøm, and S. Sorsdal, "Calculating the acoustical room response by the use of a ray tracing technique," *Journal of Sound and Vibration*, vol. 8, pp. 118–125, 1968.

[Kut14] H. Kuttruff, "Room acoustics, fifth edition," 2014. DOI: 10.1201/9781482266450. [Online]. Available: https://dx.doi.org/10.1201/9781482266450.

[MI86] P. Morse and K. Ingard, *Theoretical Acoustics*. Princeton University Press, 1986, ISBN: 9780691024011. [Online]. Available: https://books.google.nl/books?id=KIL4MV9IE5kC.

[PVG+11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[RLMS+18] D. Ramos-López, A. R. Masegosa, A. Salmerón, R. Rumí, H. Langseth, T. D. Nielsen, and A. L. Madsen, "Scalable importance sampling estimation of gaussian mixture posteriors in bayesian networks," *International Journal of Approximate Reasoning*, vol. 100, pp. 115–134, 2018, ISSN: 0888-613X. DOI: https://doi.org/10.1016/j.ijar.2018.06.004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888613X18300276.

[Sch11] D. Schröder, *Physically based real-time auralization of interactive virtual environments*. Logos Verlag Berlin GmbH, 2011, vol. 11, ISBN: 3832530312.

[SS15] L. Savioja and U. P. Svensson, "Overview of geometrical room acoustic modeling techniques," *The Journal of the Acoustical Society of America*, vol. 138, no. 2, pp. 708–730, 2015, ISSN: 0001-4966. DOI: 10.1121/1.4926438. [Online]. Available: https://dx.doi.org/10.1121/1.4926438.

[SSMS13] A. Southern, S. Siltanen, D. T. Murphy, and L. Savioja, "Room impulse response synthesis and validation using a hybrid acoustic model," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 9, pp. 1940–1952, 2013, ISSN: 1558-7924. DOI: 10.1109/TASL.2013.2263139.

[Tho17] M. R. Thomas, "Wayverb: A graphical tool for hybrid room acoustics simulation," Thesis, 2017. [Online]. Available: http://eprints.hud.ac.uk/id/eprint/33919/.

[TWH+22] P. Thoman, M. Wippler, R. Hranitzky, P. Gschwandtner, and T. Fahringer, "Multi-gpu room response simulation with hardware raytracing," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 4, 2022, ISSN: 1532-0626. DOI: 10.1002/cpe.6663. [Online]. Available: https://dx.doi.org/10.1002/cpe.6663.

[VM00] M. Vorlaender and E. Mommertz, "Definition and measurement of random-incidence scattering coefficients," *Applied Acoustics*, vol. 60, pp. 187–199, 2000. DOI: 10.1016/S0003-682X(99)00056-0.

[VS08] M. Vorlaender and J. Summers, "Auralization: Fundamentals of acoustics, modelling, simulation, algorithms, and acoustic virtual reality," *The Journal of the Acoustical Society of America*, vol. 123, p. 4028, 2008. DOI: 10.1121/1.2908264.

[Wei95a] E. W. Weisstein, *Cylindrical equidistant projection*, Web Page, 1995. [Online]. Available: https://mathworld.wolfram.com/CylindricalEquidistantProjection.html.

[Wei95b] ——, *Spherical coordinates*, Web Page, 1995. [Online]. Available: https://mathworld.wolfram.com/SphericalCoordinates.html.

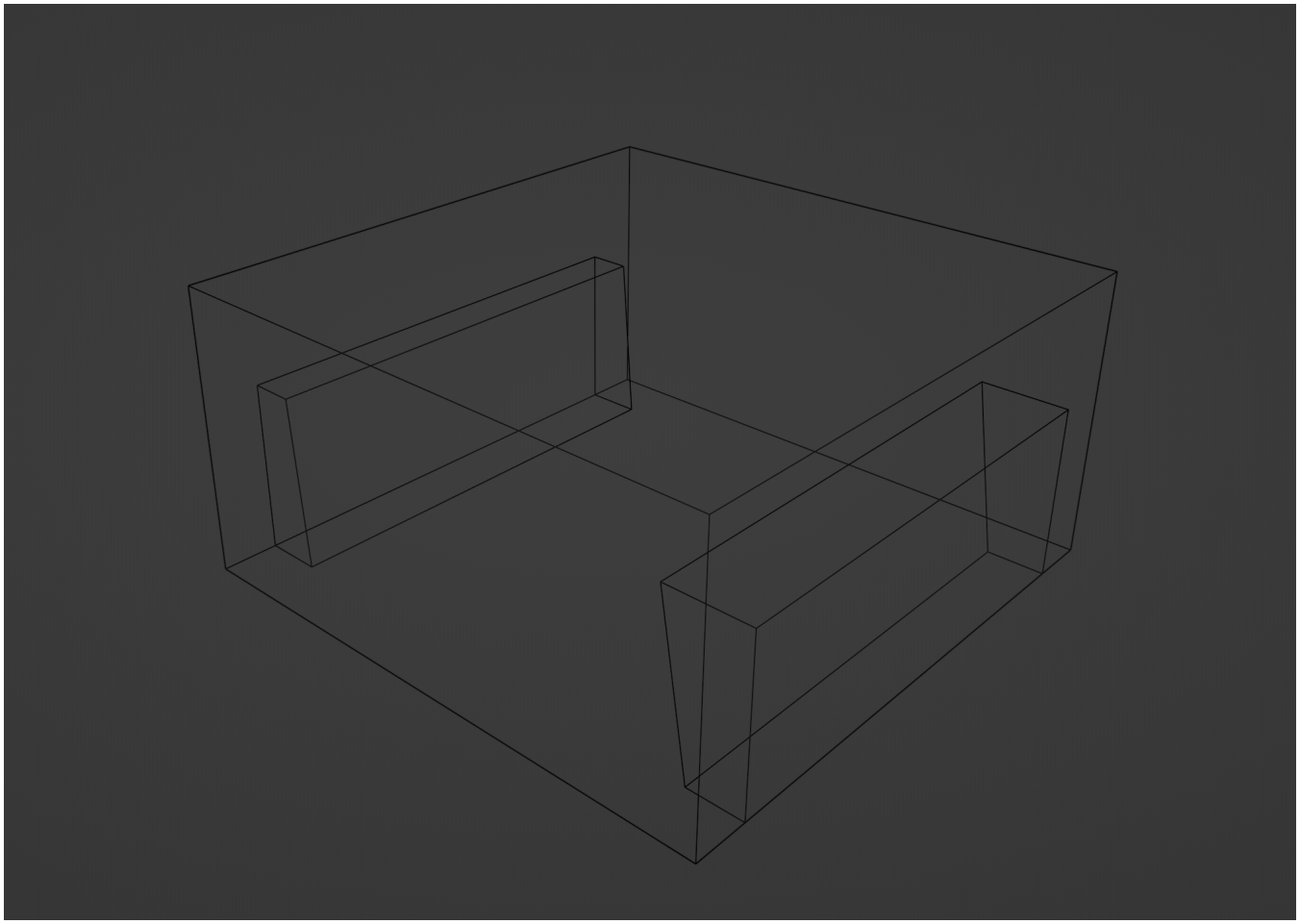


Figure 7: Room model

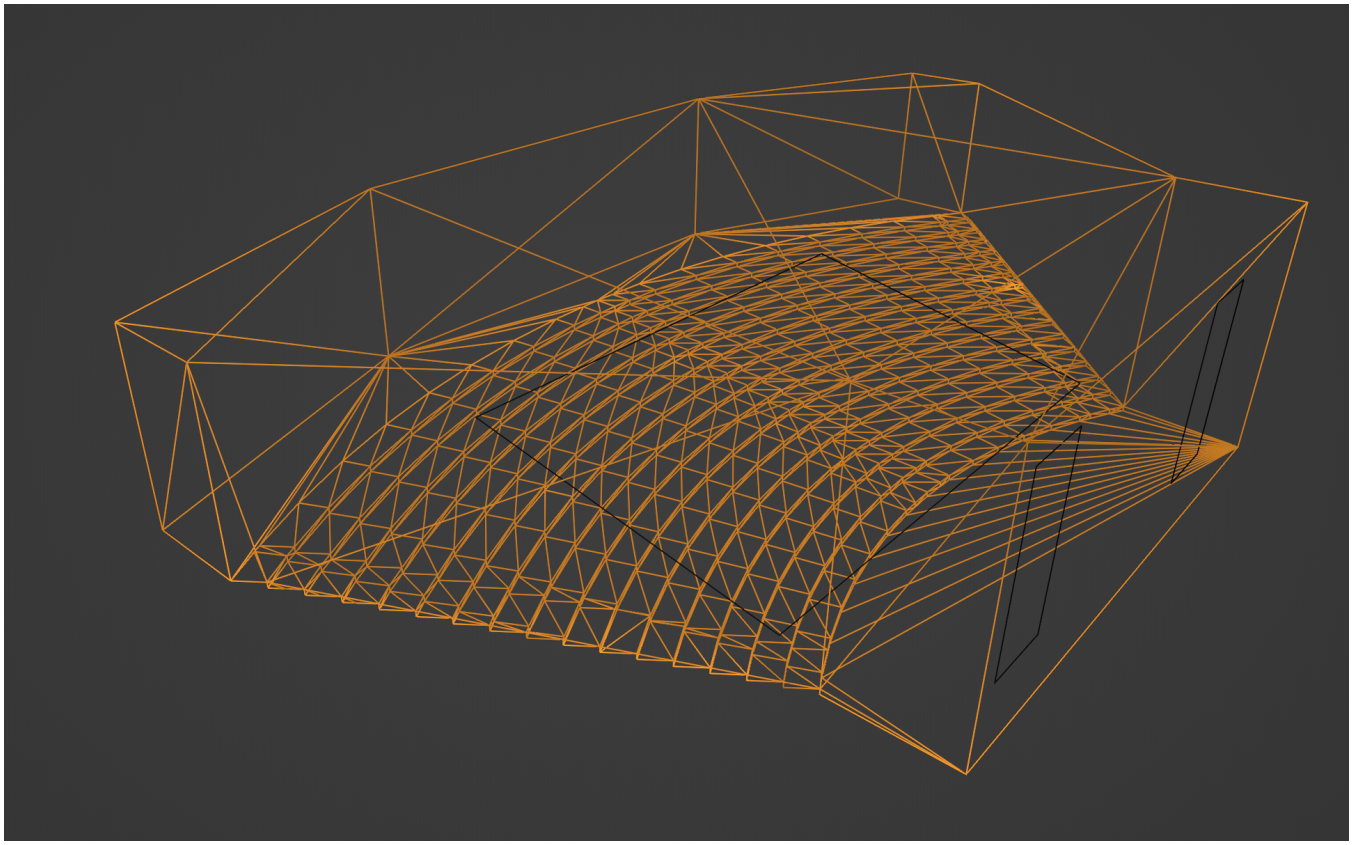


Figure 8: Theater model

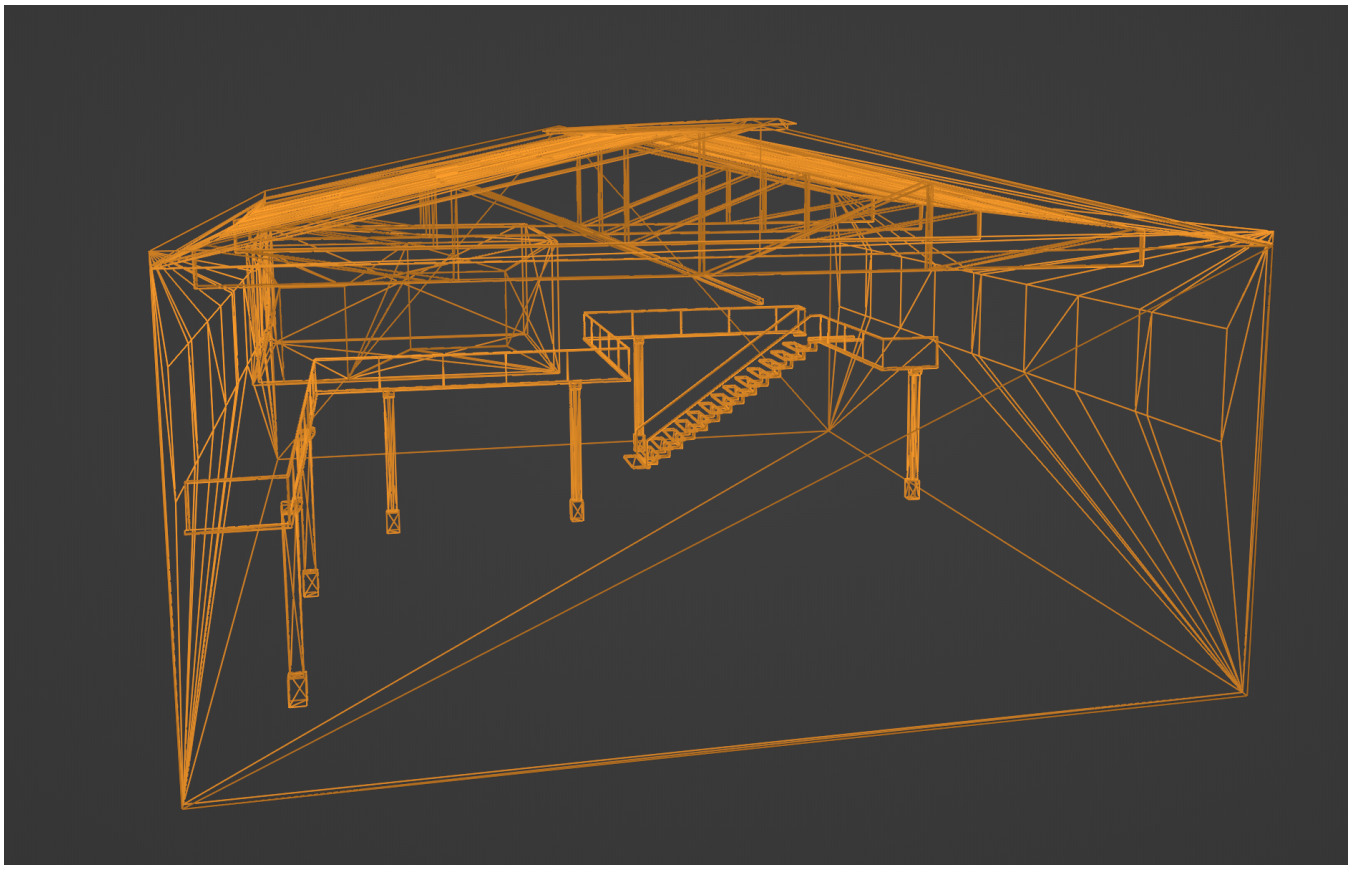


Figure 9: Warehouse model

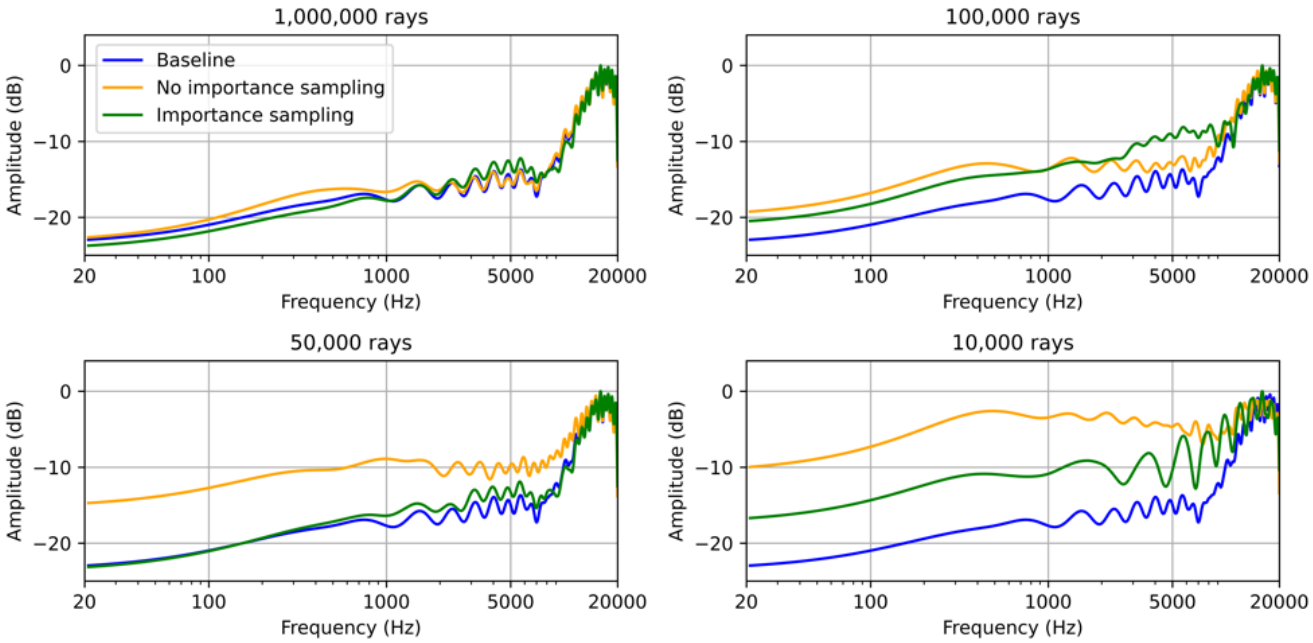


Figure 10: img/all_results/comparisons/warehouse_1_SPEC.png

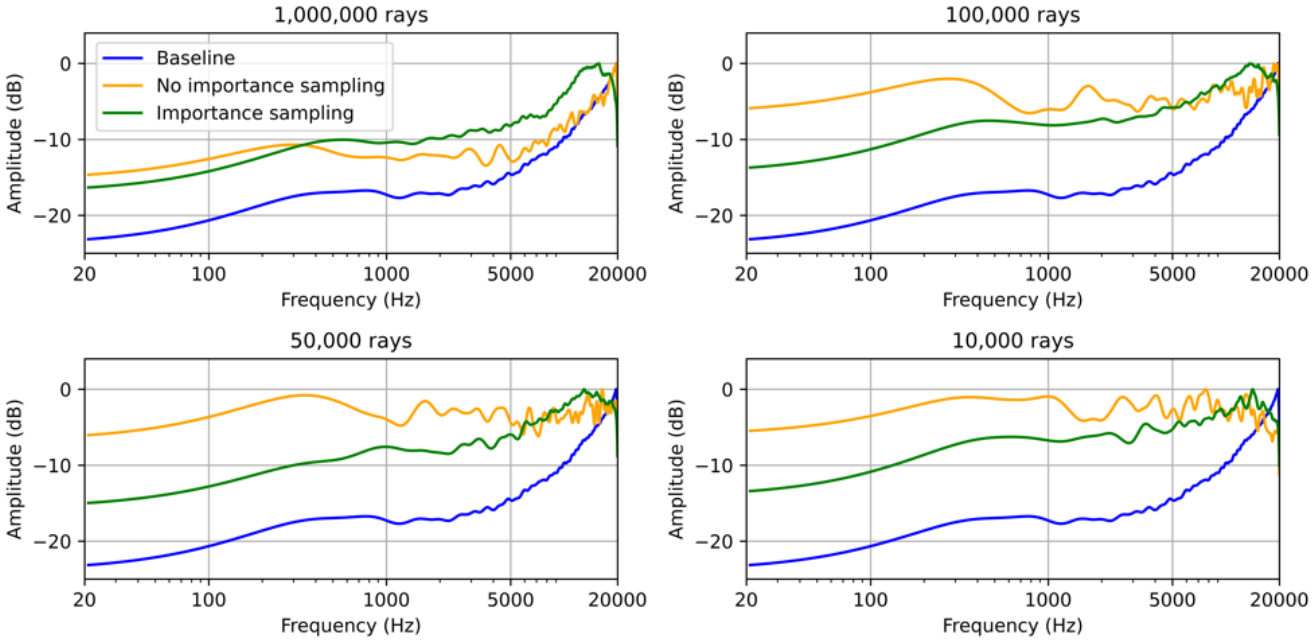


Figure 11: img/all_results/comparisons/theater_1_BOTH.png

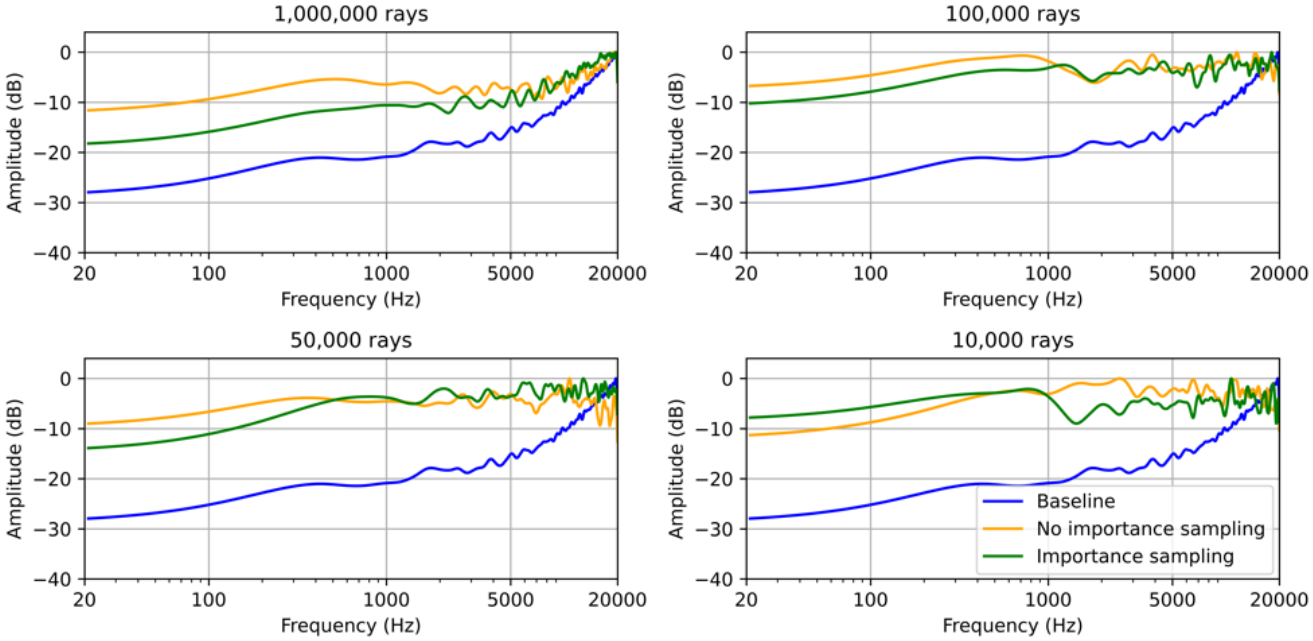


Figure 12: img/all_results/comparisons/warehouse_1_DIFF.png

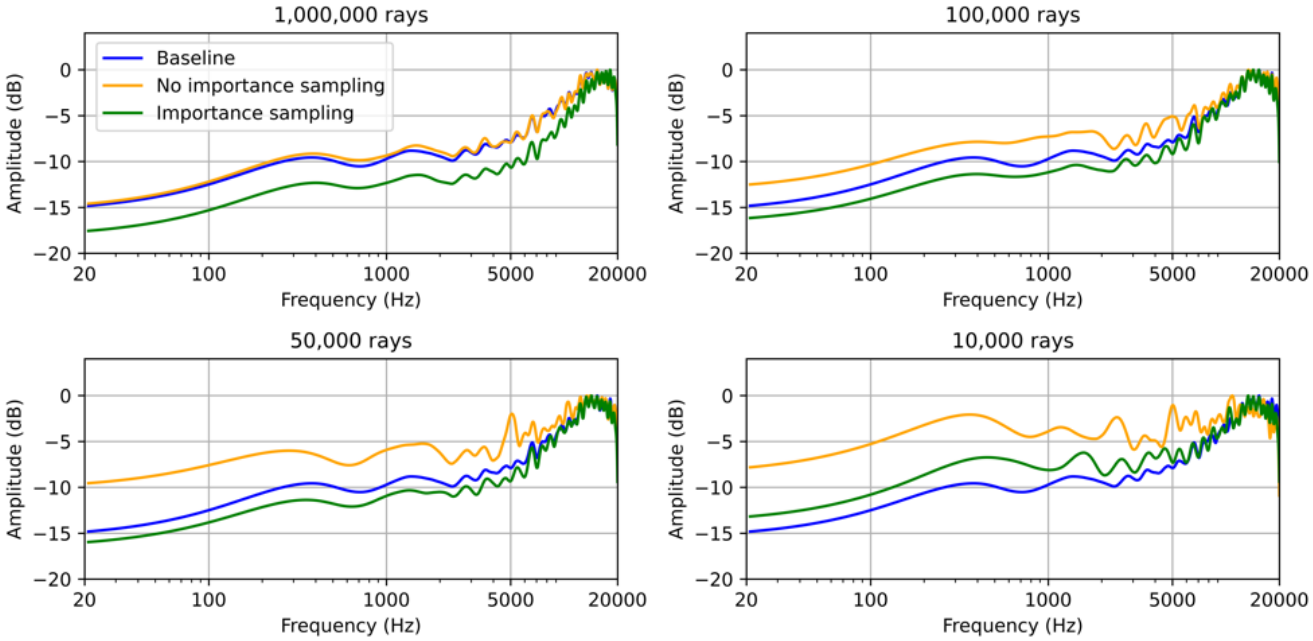


Figure 13: img/all_results/comparisons/room_1_BOTH.png

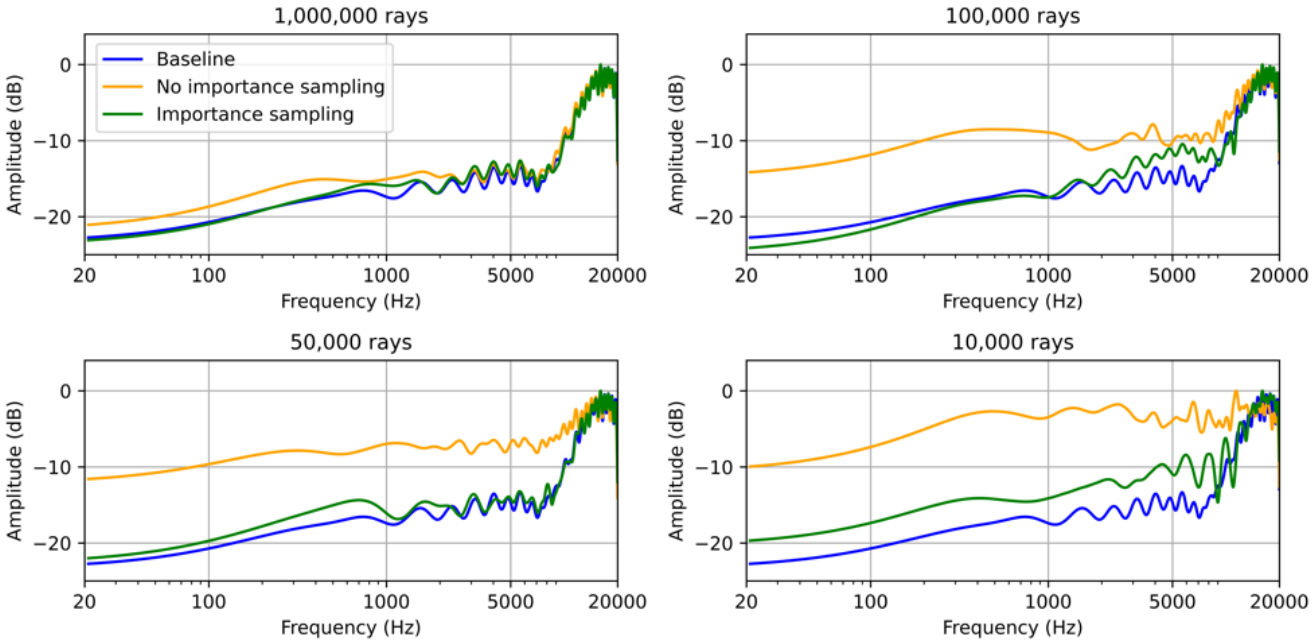


Figure 14: img/all_results/comparisons/warehouse_1_BOTH.png

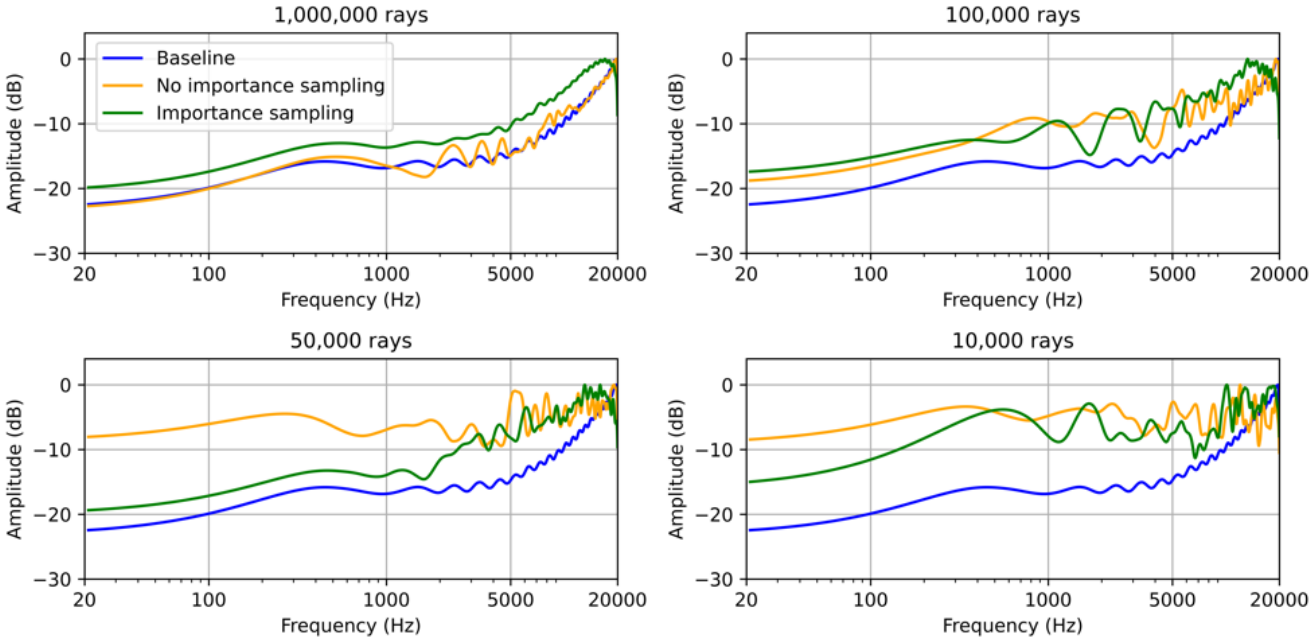


Figure 15: img/all_results/comparisons/room_1_DIFF.png

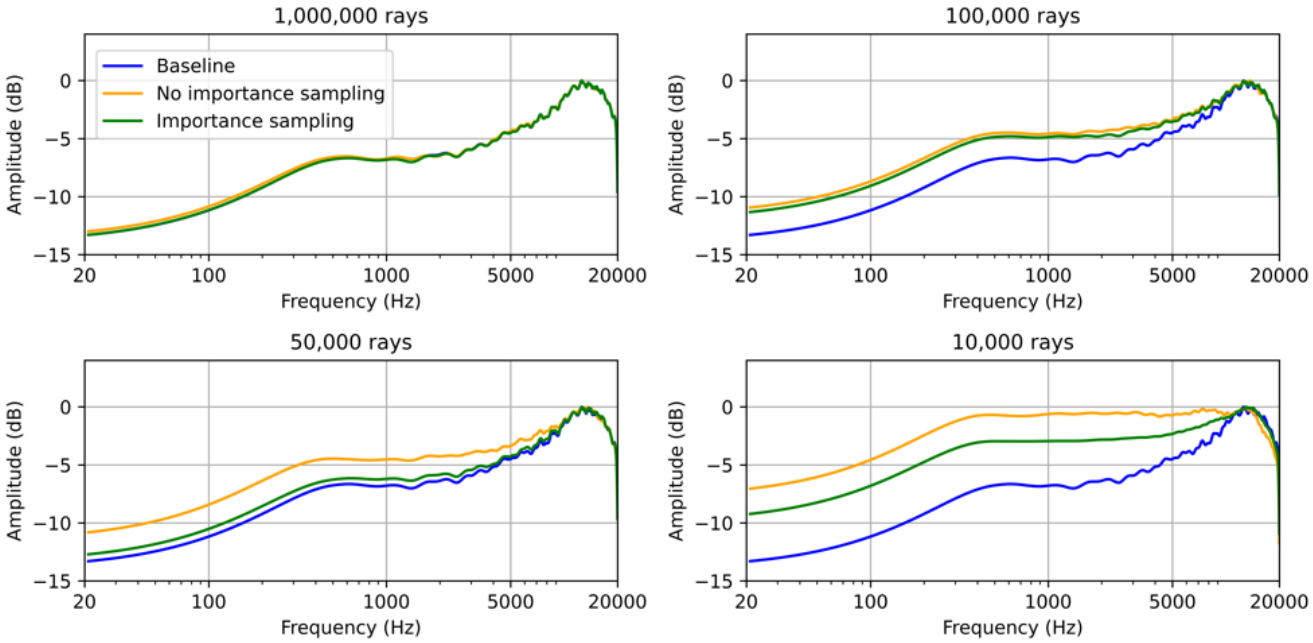


Figure 16: img/all_results/comparisons/theater_1_SPEC.png

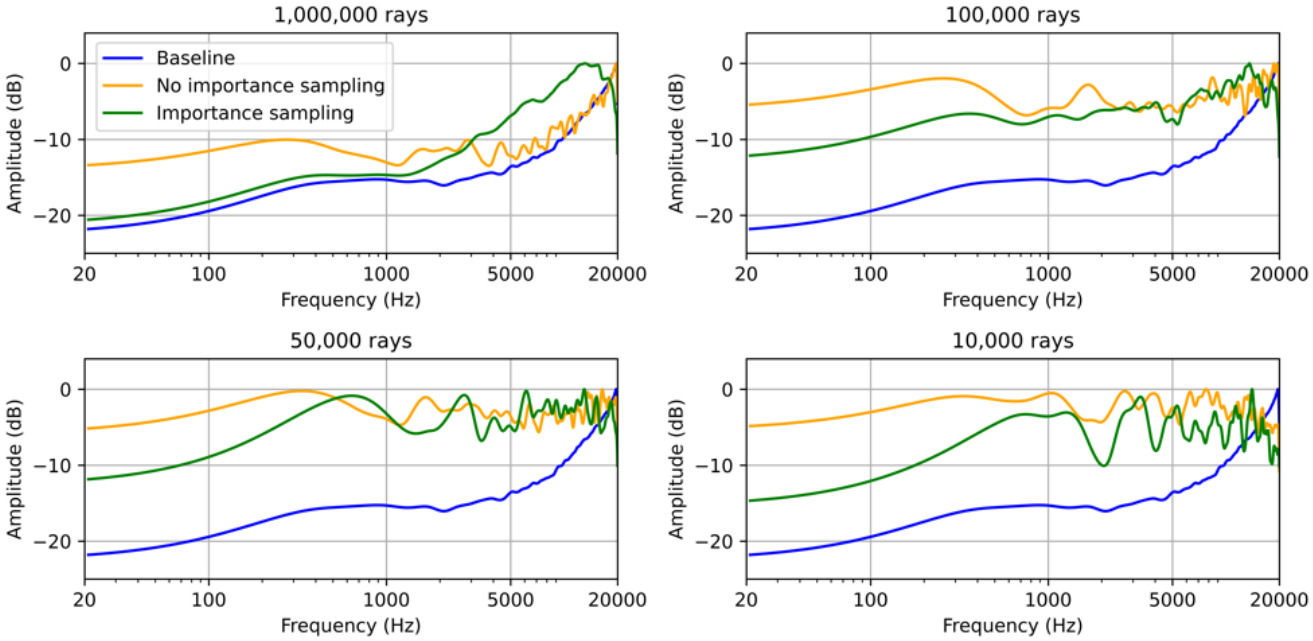


Figure 17: img/all_results/comparisons/theater_1_DIFF.png

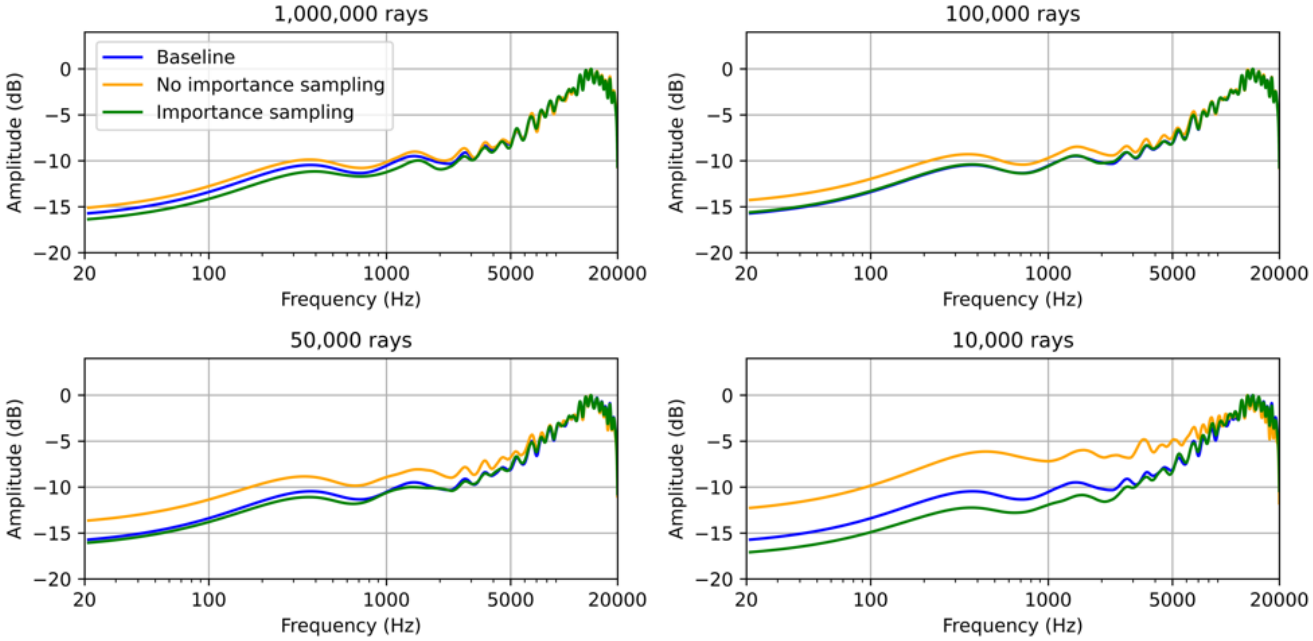


Figure 18: img/all_results/comparisons/room_1_SPEC.png

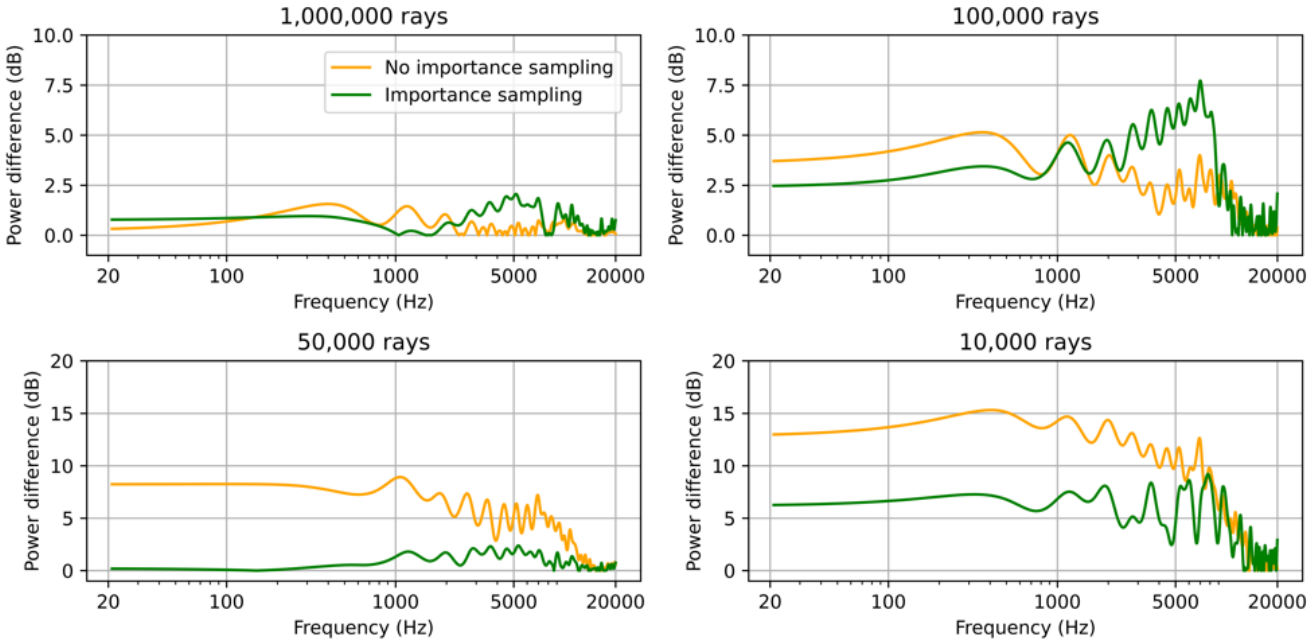


Figure 19: img/all_results/differences/warehouse_1_SPEC.png

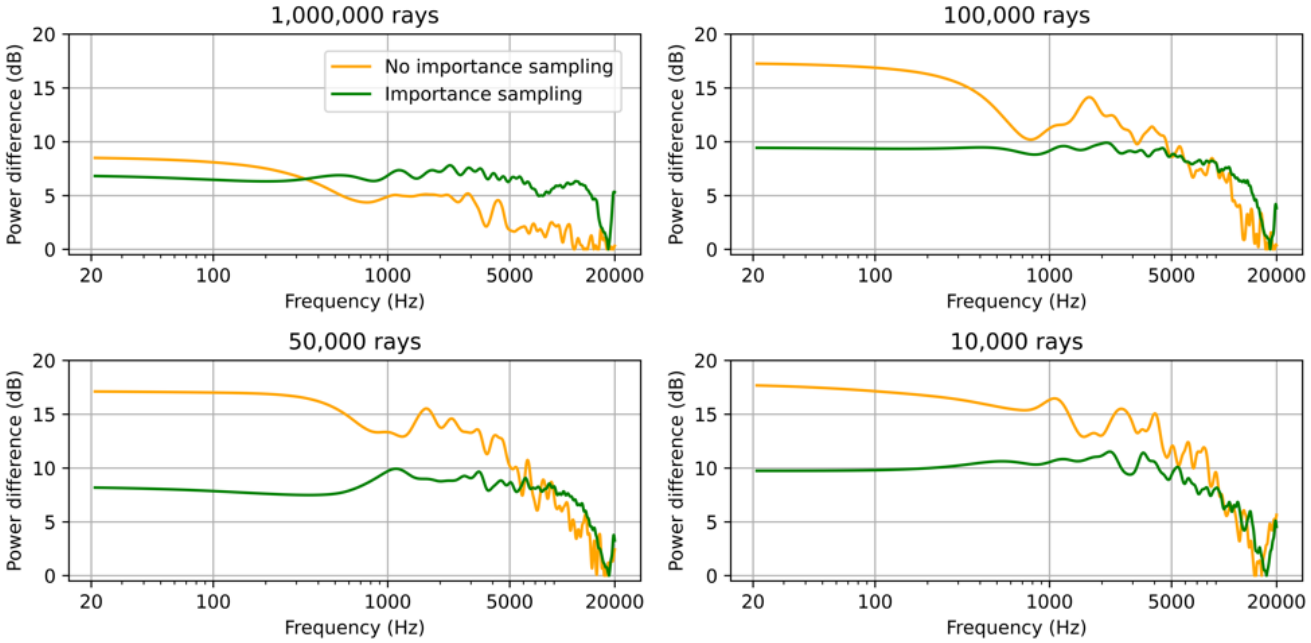


Figure 20: img/all_results/differences/theater_1_BOTH.png

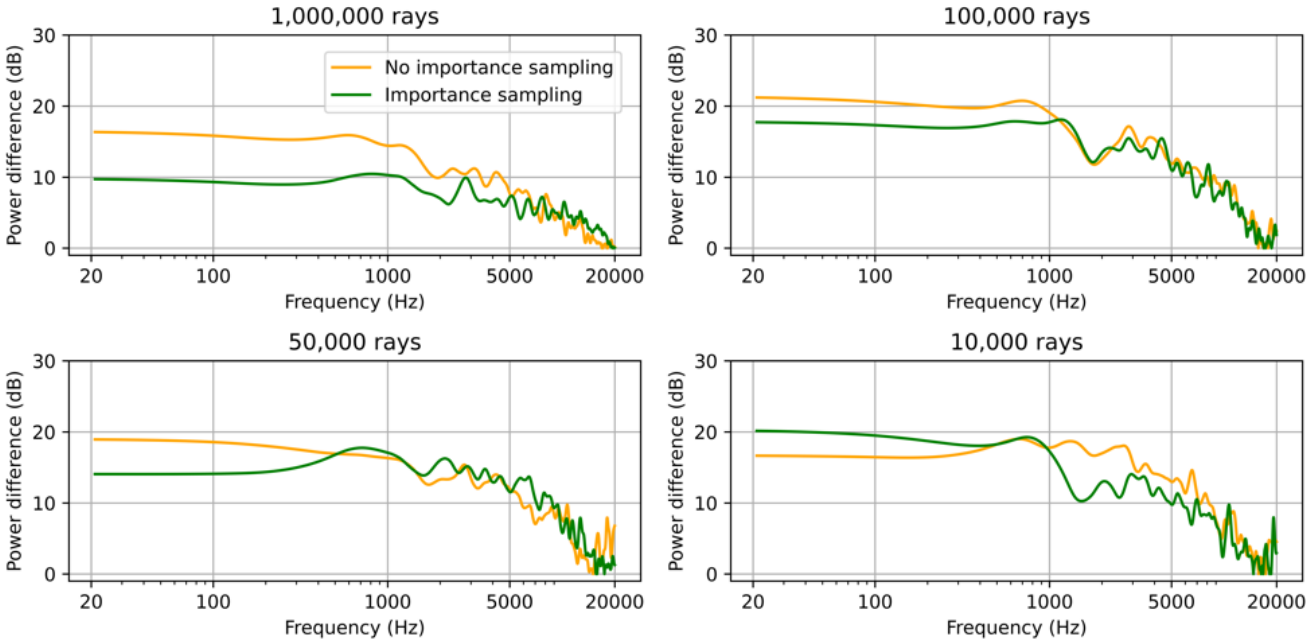


Figure 21: img/all_results/differences/warehouse_1_DIFF.png

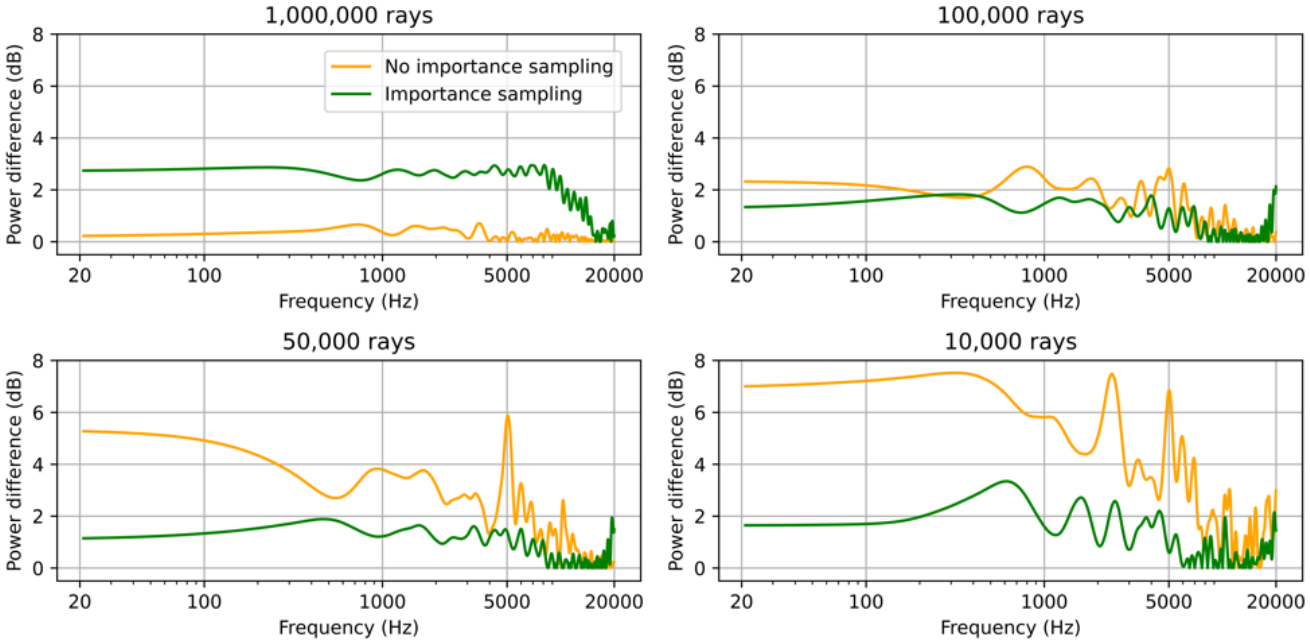


Figure 22: img/all_results/differences/room_1_BOTH.png

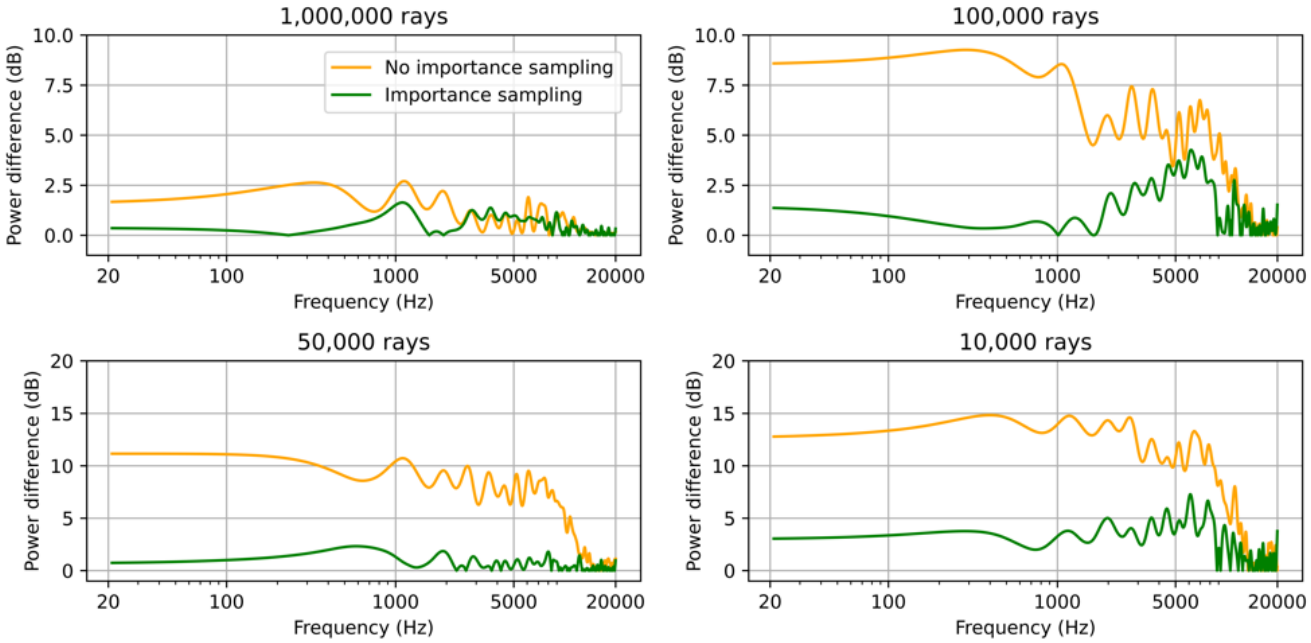


Figure 23: img/all_results/differences/warehouse_1_BOTH.png

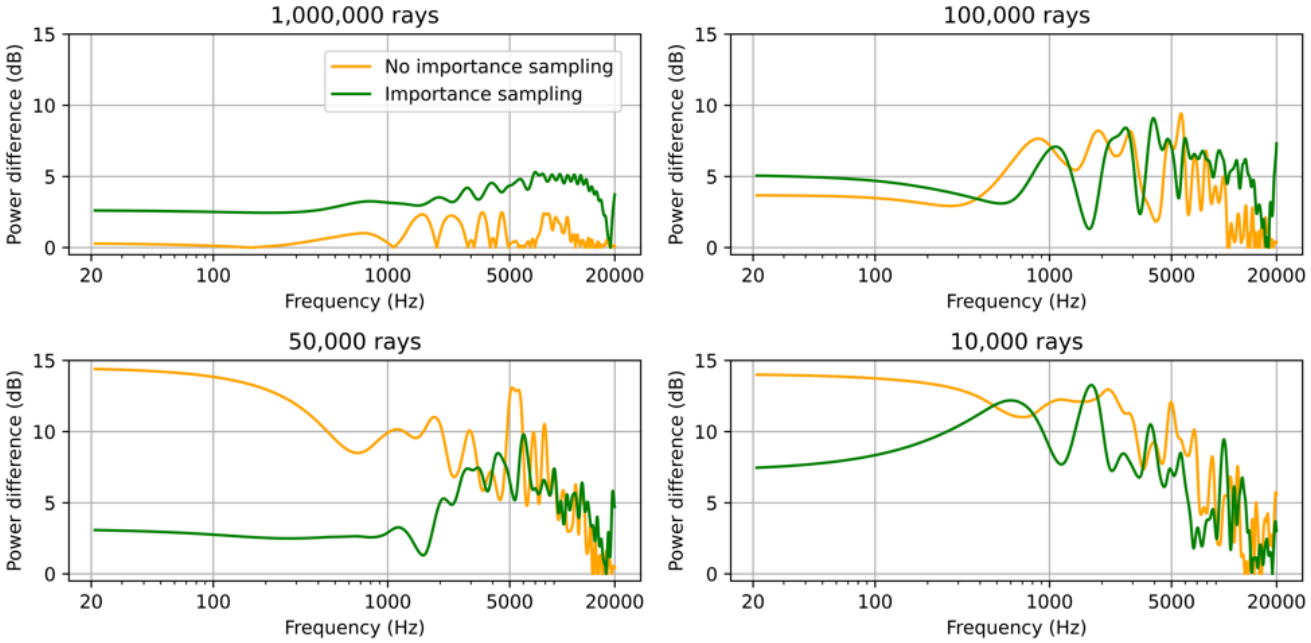


Figure 24: img/all_results/differences/room_1_DIFF.png

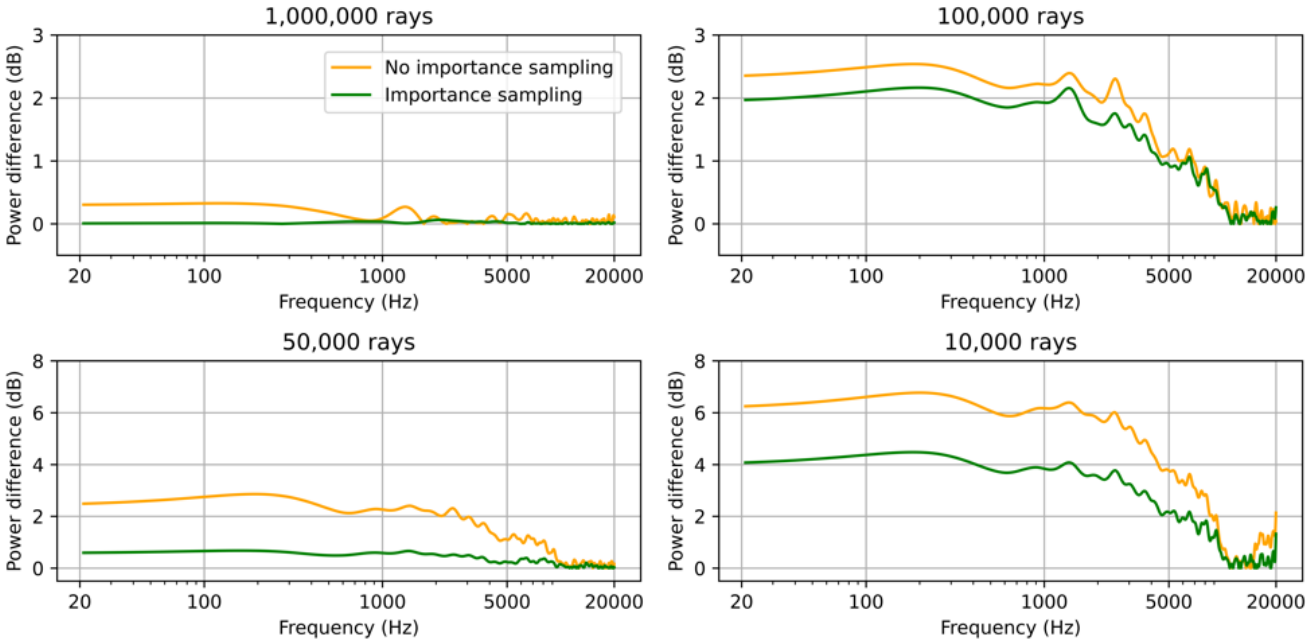


Figure 25: img/all_results/differences/theater_1_SPEC.png

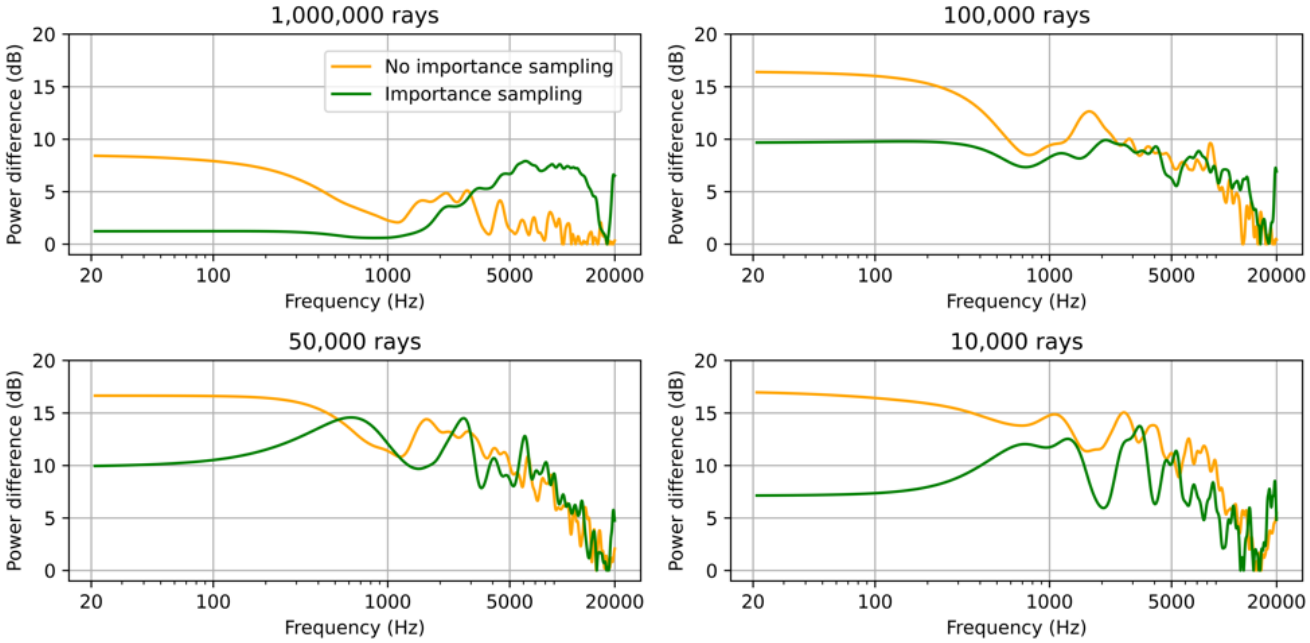


Figure 26: img/all_results/differences/theater_1_DIFF.png

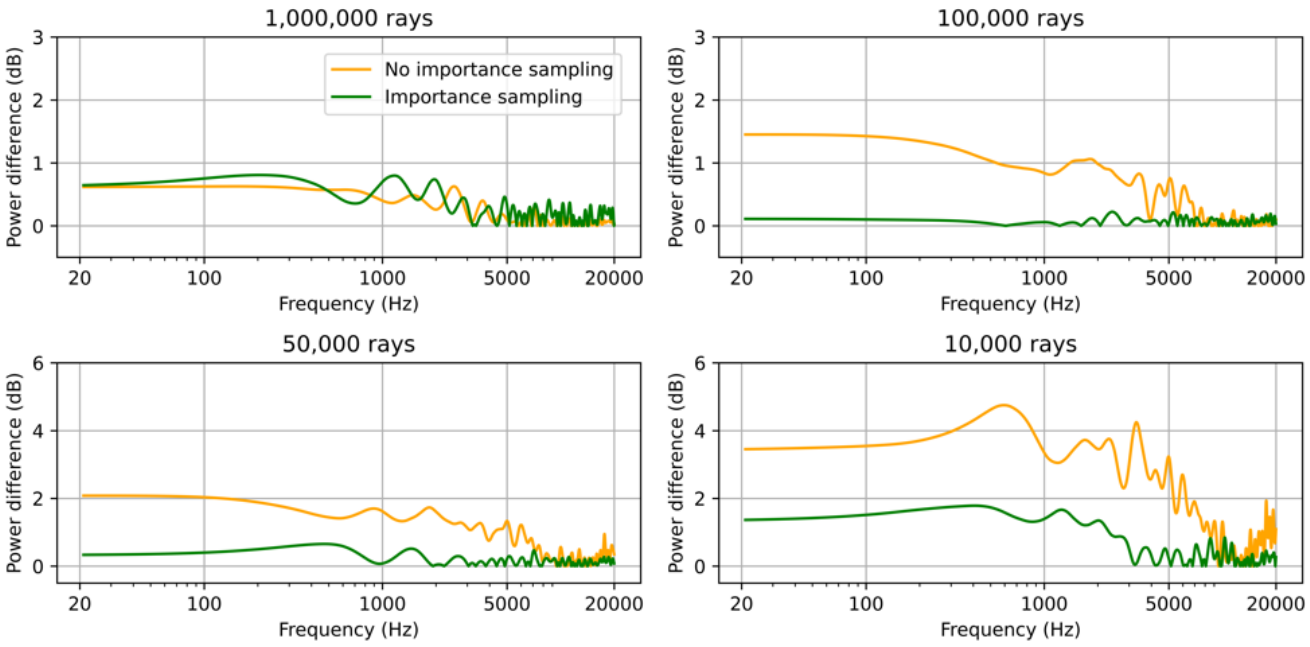


Figure 27: img/all_results/differences/room_1_SPEC.png