POLITECNICO DI TORINO

Master of Science Course in Cinema and Media Engineering

Measuring the Acoustical Properties of Recording Studios for Virtual Reality

Tutors
Prof.ssa Louena Shtrepi
Dr. Marco Fringuellino

Candidate
Camilla Tetti

27 March 2026

Table of Contents

1 Introduction

Recording studio live rooms constitute a class of acoustic spaces that differ fundamentally from traditional performance-oriented environments such as concert halls, classrooms, or auditoria. Unlike these spaces, whose acoustic design is typically driven by listener-oriented criteria and supported by consolidated target values, live rooms are primarily intended for sound acquisition through microphone systems. Consequently, their acoustic performance is not evaluated in terms of perceptual uniformity for an audience, but rather in terms of spatial sound field characteristics that influence recording quality. From an architectural standpoint, recording studios do not conform to standardized geometrical configurations. Their design is instead guided by the principles of architectural acoustics, with particular emphasis on the control of reflections, modal behavior, diffusion, and spatial energy distribution. This results in highly heterogeneous environments whose acoustic responses depend strongly on geometry, surface treatments, and intended recording use.

In this context, objective acoustic characterization becomes essential for understanding and evaluating room behavior. Standard room acoustic parameters derived from impulse response analysis, including reverberation time (RT_{60}), early decay time (EDT), clarity indices (C_{80}, C_{50}), and definition (D_{50}), provide quantitative descriptors of temporal and energetic sound field properties. While these parameters are extensively used in performance space design, their optimal ranges for recording studio live rooms are not clearly established, reflecting the functional variability of such environments.The present thesis addresses this issue through the investigation of a real recording space: the Music Lab live room at SERMIG (Turin).

The study commences with the calibration of the source in the anechoic chamber at Politecnico of Turin, a procedure that facilitates the measurement of the G level. Subsequently, an in-situ measurement campaign was initiated with the objective of extracting objective acoustic parameters from the room impulse responses. These measurements serve two functions: firstly, they characterise the existing acoustic conditions; secondly, they act as a reference dataset for subsequent modelling. The measurements conducted at SERMIG's live room encompass a comprehensive array of metrics, including the evaluation of intelligibility, the G-reference, the impulse response, and stage acoustic measurements.
A geometrical-acoustic virtual reconstruction of the space is then developed using two distinct simulation frameworks. The first of these is based on Odeon, a validated ray-tracing based acoustic simulation software. The second approach employs Blender AuVi Suite 0.7, which integrates the Python library Pyroomacoustics, thereby enabling hybrid modelling workflows that combine geometrical representation with computational acoustic processing. The simulated acoustic parameters obtained from both platforms are systematically compared with the measured data in order to assess model accuracy and ensure calibration consistency. Subsequent to validation, the virtual model is employed as a predictive instrument to investigate the impact of potential acoustic modifications.

The model facilitates objective evaluation through parameter variation and subjective assessment of

acoustic interventions via simulation and auralisation processes. This dual approach facilitates the analysis of design strategies and their influence on room response, thereby contributing to a methodology for recording studio acoustic optimisation based on measurement-driven virtual prototyping. The proposed workflow demonstrates the potential of integrated measurement and simulation techniques to support evidence-based acoustic design in environments where prescriptive target values are not formally defined.

Prior to the exploration of a specific research topic in the subsequent chapter, the acoustic design process for a recording room will be examined in greater detail. The primary case studies identified in the extant literature will also be analysed, serving as a valuable reference point for the comparison and evaluation of the measurements obtained in the case study under consideration in this thesis.

2 Live rooms acoustic design

2.1 Introduction on live rooms

Recording studio live rooms do not conform to standardized dimensional criteria and can exhibit a broad spectrum of architectural and acoustic configurations. Their volumes may range from a few cubic meters to over 5,000 m³, as exemplified by the renowned Studio One at Abbey Road. The design of each live room is inherently linked to its intended use, which varies according to the musical genre, the type of ensemble, whether an orchestra, a rock band, or a brass section, and the desired acoustic signature.

Given that these environments are purpose-built for the recording and spatial capture of the sound field through microphone techniques, their design follows the consolidated principles of architectural acoustics. Parameters such as reverberation time (T_{30}/ T_{20}), early decay time (EDT), clarity index (C_{80}/C_{50}), definition (D_{50}), and lateral energy fraction (LF) are carefully optimized to achieve the required balance, as explained below, between reverberant richness and intelligibility.

In smaller live rooms, a lower reverberation time and higher absorption are generally preferred to maintain control over direct-to-reverberant energy ratios, ensuring precision in close-miking setups typical of popular music productions. Conversely, large-scale rooms designed for orchestral or choral recordings aim to preserve a more extended reverberant field and a homogeneous spatial diffusion, enhancing the sense of envelopment and natural ambience captured by distant or spaced microphone arrays.

Therefore, the acoustic design of a live room represents a synthesis of objective parameters and subjective perception, where the physical behavior of sound within the space must support both the artistic intent and the technical requirements of modern recording practices.

Presently, the sole reference guidelines are enshrined in the UNI EN ISO 3382-1:2009 standard, entitled "Measurement of room acoustic parameters-Part 1: Performance spaces are defined as the measurement methods and typical values of the acoustic parameters that are characteristic of theatres and large concert halls.[1]

To date, international regulatory bodies have not provided guidance on the applicability of the parameters defined for live performance spaces to recording studios. In view of the marked disparity in both volume and function between the two categories of environments, it would seem expedient to consider the adoption of specific reference values that are more representative of the typical acoustic conditions of recording studios.

In the absence of specific regulations, it is still possible to refer to the parameters introduced by ISO 3382. However, it should be noted that the measured or design values may differ significantly depending on the intended use of the room and its operational purposes.

2.2 Preliminary evaluations

The design of a space intended for professional use in the field of sound recording requires a thorough preliminary assessment phase, to be carried out prior to choosing the site and defining the construction solutions.

The most significant critical factors include:

- proximity to sources of external noise, such as transport infrastructure (railways, subways, busy roads) or industrial plants, is a significant factor to be taken into consideration, since it is imperative to ensure background noise levels remain below 30 dB.
- the existence of adjacent residential structures, which are susceptible to noise emissions resulting from recording operations, has been identified as a concern. This is in accordance with the stipulations outlined in the Prime Ministerial Decree of December 5, 1997, entitled "Determination of passive acoustic requirements for buildings," and the UNI 11367:2010 standard, titled "Acoustic classification of building units." These guidelines stipulate the necessity for buildings to be designed with an apparent sound insulation rating (R'w) of at least 50 dB between distinct rooms. However, this level of insulation is inadequate in effectively mitigating the typical sound pressure levels encountered in a recording studio, which can range from 90 to 110 dB.
- The low surface mass and reduced structural rigidity of walls, floors, or roofs limit the possibility of achieving adequate sound insulation through necessary structural corrective measures. Such measures may include the insertion of mass-spring-mass systems or the creation of floating rooms (box-in-a-box). However, these measures are complex to implement and may not be economically sustainable.

In order to achieve optimal acoustic performance in a recording studio, three fundamental factors must be considered: firstly, the necessity of high structural rigidity; secondly, the importance of adequate mechanical strength of the load-bearing structures; and thirdly, the consideration of distance and shielding from external sources of disturbance.

2.2.1 Insulation and background noise

A recording studio requires high levels of insulation, both from the outside world and from other areas of the studio, such as the control room, rehearsal rooms, or other recording rooms, such as vocal booths. Furthermore, it is imperative that the equipment operates with minimal noise.

The noise in question is contingent on a number of factors, including but not limited to HVAC, acoustic infiltration, and structural vibrations. Additionally, the absence of control over low-frequency noise (i.e. noise below 125 Hz) is a salient issue that is frequently disregarded yet has a significant impact on recording systems. In high-level projects, background noise is measured in octave bands from 16 Hz to 8 kHz, with A-weighting and linear weighting, according to ISO 16283.

The quantification of this level of disturbance in spectral and perceptual terms is enabled by background noise indices. The following indices are the most useful for recording studios:

- NC - Noise Criterion (ANSI S12.2)[2], as pioneered by Beranek (1957), is a seminal index within the construction industry, having gained widespread utilisation over the years. It delineates a limit curve for the weighted sound pressure level per octave band, thereby establishing a benchmark for acceptable noise levels. Its primary application is in the United States and in HVAC systems, with typical values for recording studios ranging from 10 to 15.
- The Noise Rating (NR) system, as defined by ISO R 1996 / ISO 1996-2, constitutes a European standard analogous to NC, which was developed by the International Organization for

Standardization (ISO). The system provides a single index, designated NR, which is obtained by comparing the noise spectrum with reference curves ranging from NR 0 to NR 80. For the purpose of this study, the recommended values for recording studios are NR ≤ 20.

- The NCB (Noise Criterion, INCE, 1990) is a variant of the NC index, calibrated to engender more uniform perceptual sensations for acoustic comfort. It is utilised in American standards for high-quality sound environments. In professional settings, it is recommended to remain between values of 10 and 15.

In accordance with the reference standards and parameters delineated above for high-sensitivity recording (condenser microphones, classical music, solo vocals, mastering), background noise levels must be ≤ 20 dB(A).

2.2.2 High acoustic field diffusion

In any room there are an infinite number of available pathways for a sound wave to travel around a room and these are referred to as modes. These modes are associated to some resonances which are called eigentons, which are the natural resonance frequencies of the room.[3] The non-resonant modes are known as forced modes.

There are three types of mode in a rectangular room:

- axial modes, which exists between two parallel surfaces;
- tangential modes, which exists between four surfaces and travel parallel to the other two;
- oblique modes, whose paths involve striking all six surfaces before returning to their point of origin to begin retracing the paths.

The configuration of a room is the primary factor influencing its modal patterns. When the sound source is relocated within the environment, the distribution of the acoustic energy produced by the room modes changes; however, the modal pattern itself remains unaltered. In order to achieve greater control over these phenomena, spaces designed for high-quality audio recording are often constructed with non-parallel walls.[3]

This design choice offers an additional advantage in that it helps to prevent flutter echoes too.

The management of early reflections is equally important, as they should arrive within approximately 20 milliseconds to reinforce and enhance the direct sound, taking advantage of the Haas effect. In order to achieve this objective, it is essential that the room ensures good lateral efficiency, thereby enhancing the perceived spaciousness of the sound field. This necessitates the predominance of lateral reflections over those from the ceiling; consequently, a high ceiling is typically favoured.[4]

A pragmatic approach to mitigating these reflections entails the deployment of diffusers, which serve to redirect sound energy towards the side walls. Alternatively, in the event of a ceiling that is too low, the use of sound-absorbing materials can be employed in order to minimise unwanted reflections.

The regulation of sound sources within such environments thus necessitates meticulous consideration of sound field diffusion, a factor that plays a pivotal role in psychoacoustic perception. Diffusion has been shown to enhance listening comfort by reducing the listener's ability to localise sound sources or perceive

dominant directions. Furthermore, diffusion has been demonstrated to enrich the sense of spatiality, giving the impression of being in a larger and more enveloping space.
The implementation of acoustic diffusers is instrumental in achieving this objective. The purpose of the elements in question is to disperse sound waves in a manner consistent with the desired acoustic environment. The configuration of these elements is intended to disperse sound waves in multiple directions, thereby ensuring a uniform distribution of acoustic energy throughout the room. Consequently, the sound field becomes more diffuse and homogeneous, thereby reducing unwanted energy concentrations and excessive direct reflections.

2.2.3 Key Acoustic Parameters for the Recording Studio evaluation

The acoustic quality of music recording studios has been widely acknowledged as a fundamental factor influencing both the performance of musicians and the effectiveness of recording processes. However, in contrast to the case of large performance venues, the acoustic characterisation of recording environments is not yet subject to universally accepted standards. As has been emphasised in recent literature, including the work of von Lauer-Münchhofen et al., current guidelines are largely dependent on best-practice approaches as opposed to target values which are strictly defined.[5] This absence of standardisation signifies that it is not possible to describe the acoustic characteristics of a recording studio by means of a single metric. Instead, the acoustic behaviour of these systems must be evaluated through a combination of energetic, temporal, spectral, and geometric parameters derived from the room impulse response.

2.2.3.1 Objective acoustic parameters

The majority of recent studies adopt measurement frameworks derived from ISO 3382 methodologies, adapting them to the specific needs of recording spaces. In particular, the study by von Lauer-Münchhofen et al. analysed 20 recording rooms used for pop, rock, and jazz music, extracting a set of core parameters from measured impulse responses.[5] These includes the following:

- Reverberation time (T_{20}, T_{30})
- Early Ddecay Time (EDT)
- Clarity (C_{80})
- Centre time (T_s)
- Sound strenght (G)

In a similar vein, the investigation of a live recording room by Tavelidou et al. corroborates the pertinence of the aforementioned family of parameters, underscoring their perceptual interpretation. The findings of the study demonstrated a positive correlation between the perceived "liveness" of the space and the reverberation time. The results also indicated that the Early Decay Time (EDT) exhibited a stronger alignment with subjective impressions of reverberation.[6] Energy-based parameters, such as C_{80} and D_{50}, have been shown to describe the relationship between early and late reflections. This relationship provides information about musical clarity and intelligibility. It is evident that low values of centre time Ts are

indicative of a pronounced concentration of energy in the early segment of the impulse response. This phenomenon is concomitant with enhanced clarity and control during the process of recording.[6] These findings serve to reinforce the notion that the concept of reverberation time alone is insufficient for the description of recording spaces. As explicitly stated in von Lauer-Münchhofen et al., future guidelines should incorporate additional parameters related to reflections, spatial perception, and timbral characteristics.[5] In addition to objective measurements, interviews conducted with recording engineers reveal a consistent set of perceptual descriptors used in practice. Engineers frequently characterise recording rooms in terms of reverberation, spatial impression, colouration, early refelctions and room resonances. [5]
This highlights a clear relationship between measurable acoustic indicators and subjective perception. It also confirms that recording environments must balance technical control with musical expressiveness.

2.2.3.2 Acoustic behaviour of recording rooms

It is evident from a comprehensive review of the extant literature that recording rooms can be broadly classified into three distinct acoustic types. Neutral or "dry" rooms have been shown to exhibit very short reverberation times, with mean T_{30} values of approximately 0.19 s. These are typically associated with high acoustic control and versatility. Conversely, live rooms exhibit prolonged reverberation times, approximately 0.40 s, thereby engendering a more pronounced spatial impression and providing enhanced natural acoustic support for instruments such as drums and vocals. A third category comprises so-called coloured rooms, distinguished by excessive low-frequency reverberation, which can reach up to approximately 0.63 s and frequently gives rise to an acoustically imbalanced response.
The observations presented herein suggest that the optimal conditions for desirable reverberation in the context of recording studios are situated between those of highly controlled listening environments and more reverberant rehearsal spaces. This is characterised by a balance between clarity and acoustic support. In practical applications such as pop and rock production, the duration of reverberation is typically maintained below one second, as excessive reverberation is difficult to remove once recorded, but can be added during post-production.[5]

2.2.3.3 Geometrical consideration

In addition to energetic parameters, the physical configuration of recording environments strongly influences acoustic performance. According to architectural acoustics guidelines summarised by Spagnolo (2015), control rooms should typically exhibit: volume ≥ 80 m³ and a reccomended height between 3 m and 4 m. [7] This to have a proper modal distribution, avoiding standing wave accumulation.
The relationship between optimal reverberation time and room volume for different types of performance spaces follows the classical design curves proposed by Beranek (1962), which are widely adopted in architectural acoustics literature.

2.2.3.4 Synthesis

Table 2.1 Comparative overview of recording and rehearsal spaces investigated in the reviewed literature, including room type, reverberation characteristics and key acoustic descriptors.

Study	**Authors/year**	**Environment Analysed**	**Room type**	**Volume $[m^3]$**	**Acoustic parameters used**	**Key findings**
Applicability of Common Room Acoustic Parameters for Music Recording Spaces	von Lauer-Münchhofen et al., 2025	S1 R1	Type I - T_{30} = 0.29 s	98	T_{30}, C_{80}, STearly, G, BR, TR	
		S1 R2	Type I - 0.16 s	20		
		S2 R1	Type II - 0.38 s	88		
		S2 R2	Type I - 0.29 s	64		
		S2 R3	Type II - 0.41 s	62		
		S3 R3	Type II - 0.36 s	73		
		I1 R1	Type I - 0.18 s	59		
		I1 R2	Type I - 0.17 s	40		
		I1 R3	Type I - 0.17 s	24		
		I2 R1	Type I - 0.21 s	88		
		I2 R2	Type I - 0.18 s	56		
		I3 R1	Type III - 0.51 s	136		
		I3 R2	Type III - 0.60 s	136		
		I3 R3	Type III - 0.28 s	30		
		I4 R1	Type I - 0.19 s	116		
		I4 R2	Type I - 0.17 s	47		
		I5 R1	Type II - 0.49 s	645		
		I5 R2	Type I - 0.13 s	27		
		I6 R1	Type III - 1.14s	666		
		I6 R2	Type II - 0.34 s	455		
Live Room Acoustic	Tavelidou et al., 2016	Sound studio live room (University of Athens)	Live recording room	*	T_{30}, EDT, C_{50}, C_{80}, D_{50}, Ts, IACC	Low T_S and high clarity indicate strong early energy concentration

Evaluation Methods of the Acoustic Quality of a Small Recording Studio	Leccese et al., 2015	Small recording studio ("Il Musicante", Italy)	Recording studio	100.89	Modal behaviour, frequency regions, RT	Acoustic performance is strongly linked to modal distribution
Acoustic assessment of four music rehearsal rooms	Şaher et al., 2020	Rehearsal Room 1	Rehearsal space RTmid within recommended range (0.92 s)	233	RTmid, frequency balance	
		Rehearsal Room 2	Rehearsal space RTmid slightly above range (0.74 s)	192	RTmid, frequency balance	
		Rehearsal Room 3	Rehearsal space RTmid close to target values (0.63 s)	163	RTmid, frequency balance	
		Rehearsal Room 4	Rehearsal space RTmid outside range (0.98 s)	292	RTmid, frequency balance	

In relation to the live room designed by Tavelidou et al. (2016), the publicly available documentation from the Laboratory of Music Acoustics and Technology at the University of Athens describes the studio as a structure consisting of several rooms, with a total surface area of approximately 120 m². However, the specific geometric characteristics of the live room analysed, such as volume or proportions, are not reported.

The study by von Lauer-Münchhofen et al., 2025, analysed the recording rooms of three professional studios, identified as "S", and seven universities or educational institutions, identified as "I". The paper explicitly states that the respondents were often members of staff at the locations measured and therefore a methodological choice was made to anonymise them, which is a fairly typical practice in acoustic studies. Complessively in this study it is shown how Recording rooms fall between control rooms and rehearsal rooms for their acoustic behaviour. Furthermore, the categorisation of the space was undertaken in accordance with the resulting T_{30}, as illustrated in the accompanying Table 2.2.

Table 2.2 cluster summary for von Lauer-Münchhofen classifications of live rooms

Type	**Description**	**Mean T_{30}**
Type I	Neutral or dry rooms	~ 0.19 s
Type II	Live rooms	~ 0.40 s
Type III	Coloured rooms	~ 0.63 s

On the other hand, the study by Şaher et al. (2020) found that the perceived acoustic quality of the environments analyzed is closely related to RTmid values compared to the ranges recommended by ISO 2359 1 for music rehearsal rooms. In particular, for spaces intended for high-powered instruments, such as drums or amplified instruments, the recommended average reverberation time is between approximately 0.55 and 0.80 s, while for environments intended for weaker acoustic instruments (such as strings or wind instruments), the suggested range is between 0.75 and 1.00 s. Rooms with RTmid values within these ranges were generally preferred by musicians.Although reverberation measurements were conducted according to the procedures set out in ISO 3382 [8], Şaher et al. report the results in terms of RT_{mid} rather than individual T_{30} values, as ISO 2359 1 evaluates the acoustic quality of rehearsal rooms using the average reverberation time in the mid frequencies.[9,10]

The RT_{mid} parameter is not a separate calculation method from T_{30}, but rather the average of the reverberation times, generally derived from T_{30} values, in the mid-frequency bands, typically at 500 Hz and 1 kHz.

While the studies by von Lauer-Münchhofen et al., Şaher et al., Leccese et al., 2015 and Tavelidou et al., 2016 provide case-based acoustic measurements, additional literature offers complementary perspectives. In particular, UTET emphasises the pivotal role of geometrical and modal properties in determining that proper room proportion enhance a better acoustic behaviour perception of a room[7] , while the ACE Challenge confirms the relevance of reverberation-related parameters as key descriptors of enclosed spaces.[11]

Table 2.3 Literature-based observed ranges of selected acoustic parameters for recording spaces.

Parameter	**Observed range**	**Interpretation**
T_{30}	0.19 - 0.40 s [12]	A range in which the space won't be too dry nor too *wet*
EDT	~0.15 - 0.45 s [6]	Closely linked to perceived reverberation
C_{80}	10 - 30 dB [5]	High values means high perception of control of the sound
T_s	~40 e 70 ms [5]	If low it indicates strong early energy
G	15 – 25 dB [5]	Suitable acoustic support

The valued in Table 2.3 are inherently frequency-dependent, as the early decay is strongly influenced by the distribution of early reflections and modal behaviour across the spectrum; consequently, low-

frequency energy build-up or mid-high frequency control may lead to perceptually different impressions even when overall EDT values appear similar. [5,7]

With regard to the recommended range of T_s values, shown in Table 2.2, the interval between 40 ms and 70 ms facilitates the delineation of a type 1 space, as defined by the von Lauer-Münchhofen classification, which is characterised by its high level of control. Indeed, a low centre time typically signifies that energy is concentrated in the early reflections, thereby engendering greater clarity, an outcome that is particularly advantageous in a recording studio of modest proportions.

In most cases, the same recording room is used for recording very different musical genres and in different logistical situations. This means that variable acoustics are required, using curtains, elements that can be opened or adjustable surfaces, which can alter the reverberation time as needed.

3 Theoretical Background

3.1 Key Acoustic Parameters

3.1.1 Reverberation time

Reverberation time is a fundamental indicator for evaluating the acoustic characteristics of an environment. The term "time constant" is defined as the time interval required for the sound decay curve, obtained from the impulse response of the system, to reduce by 60 decibels (dB) compared to the initial level after the sound source has been interrupted.
In scenarios where background noise levels restrict the feasible range of signal dynamics to below 30 dB for each frequency band, the adoption of shorter decay intervals, such as T_{20} or T_{30}, becomes a viable option. By employing linear extrapolation techniques, the T_{60} value can be derived from these measurements.
The descriptor T_{60} can be measured with the Sabine's formula, which can be applied to relatively large rooms, not too deaf and with walls that create diffuse reflections. Hence having the volume ($V\,[m^3\,]$) of the ambient and the total absorption ($A_{tot}\,[m^2]$) with c = 343 ms e T_{air}= 20°C, the formula 4.1.

$$T_{60} = 0.161 \frac{V}{A_{tot}} \text{ (s)} \qquad (4.1)$$

It can be defined an optimal value for the T_{60} depending on the room volume and on its intended use. Usually the T_{60} obtimal values are referred to the mid-high frequency bands 500 Hz or 1kHz, but for a better acoustic quality low frequencies are not omitted from the assessment.
The subject of this thesis, with a volume of 105.66 m^3 and intended for use as a recording hall for various genres, currently has a reverberation time of 0.57 seconds at a frequency of 1 kHz.
The international standard EN ISO 3382 (last revised in 2009), entitled "Measurement of Reverberation Time of Rooms with Reference to Other Acoustic Parameters" [13], establishes the protocol for measuring reverberation time and delineates two primary methodologies. The first method is the interrupted noise method. The second method is the impulse response method, which is based on the analysis of the energy decay of the impulse response of the environment

3.1.2 Early Decay Time (EDT)

It has been demonstrated that, particularly within the domain of music, the decay intervals indicated by traditional reverberation times are rarely perceivable. Research conducted by Atal, Schroeder, and Sessler has demonstrated that the human auditory system predominantly perceives the initial segment of the decay curve, specifically the first 160 milliseconds, a phenomenon referred to as "running reverberation." In 1970, Jordan proposed the Early Decay Time (EDT) parameter, which measures the decay rate of the first 10 decibels (dB) of the curve. [14]

The early reflections' intensity and level have a considerable impact on EDT, which, in contrast to traditional reverberation time, can exhibit significant variations depending on the position of the source or receiver within the same environment. This renders it a particularly suitable descriptor for representing the subjective attribute of reverberation.
Given the established correlation between subjective perceptions of reverberation and Early Diffusion Time (EDT) rather than Reverberation Time (RT), design decisions should prioritize this parameter. However, it is imperative to assess the perceptibility of design modifications affecting EDT to the listener, as indicated by Just Noticeable Differences (JND) [14].

3.1.3 Definition (D_{50})

$D_{50}$is a temporal acoustic descriptor that quantifies the ratio between the sound energy incident on the listener within the first 50 milliseconds of emission and the total sound energy. The introduction and formalisation of this parameter, which aims to objectively characterise the clarity of the acoustic signal, can be traced back to the research conducted by Meyer and Thiele. They defined this parameter as "definition" (originally Deutlichkeit)[3] , mathematically expressed in equation 4.2, and subsequent contributions by Thiele (1953) and Haas (1951; 1972) on the so-called "law of the first wavefront" further developed this concept.
This phenomenon, termed the "precedence effect" or "Haas effect," postulates a psychoacoustic mechanism wherein the perceived location of a sound source is predominantly influenced by the initial wavefront that reaches the listener. Subsequent reflected components do not modify directional perception, provided that their time delay remains below a critical threshold. Beyond this threshold, they are perceived as distinct events or as a degradation of sound definition.

$$D_{50}=\left[\frac{\int_0^{50ms}[g(t)]^2dt}{\int_0^{\infty}[g(t)]^2dt}\right]\cdot 100 \qquad (4.2)$$

From equation 4.2 we can clearly see how both integrals must include the direc sound and D will be 100% if the impulse response does not contain any components with delays in excess of 50 ms.
The definition is expressed in percentage values. It is evident that an increase in the magnitude of the sound, indicated by a higher percentage value (e.g. 50%), corresponds to a heightened perception of the auditory stimuli.
It is closely linked to other parameters, such as C_{50} (Clarity Index), but with the different intent of characterising the transparency of music in a concert hall.[3]

3.1.4 Clarity index (C_{50} , C_{80})

The clarity index, despite the possibility of erroneous formal equivalency with the definition, aims to delineate the transparency of music in a concert hall, a concept introduced by Reichard. The objective of

this descriptor is to quantify the balance between sound energy that arrives early and that which arrives late.
The decibel is expressed in decibels and is mathematically defined as in equation 4.3.

$$C_t = 10\log_n \left(\frac{\int_0^t p^2(t)dt}{\int_t^\infty p^2(t)dt} \right) \ [dB] \tag{4.3}$$

It is posited that p(t) denotes the sound pressure of the impulse response, and that t is the integration time threshold which separates early and late reflections. Indeed, modifying the integration time threshold results in a change in the auditory characteristics. In order to ascertain the speech clarity the index C_{50}, with a time threshold of 50 milliseconds is employed. Conversely, in the context of music, which is known to benefit from augmented early reflection, the index C_{80} is measured over the initial 80 milliseconds.

3.1.5 Energy centre time T_s

The instantaneous energy centre of gravity T_s, as delineated by ISO 3382, denotes the temporal centre of gravity of the sound energy contained within the impulse response. Introduced by Kuerer (1969), this parameter has been demonstrated to exhibit a robust correlation with the perception of clarity, both in speech and music.
From a perceptual standpoint, low values of T_s indicates that the majority of the energy is concentrated in the early reflections, thereby affording the sound enhanced definition and intelligibility. Conversely, elevated values are indicative of a predominance of late energy, which is associated with a perception of reduced clarity and a sensation of heightened reverberation. [15]
In summary, the barycentric instant of energy can be interpreted as the weighted average time over which sound energy is distributed, and is formally expressed through the equation 4.4.

$$T_s = \ 10\log_{10} \left(\frac{\int_0^\infty t \cdot p^2(t)dt}{\int_0^\infty p^2(t)dt} \right) \ [ms] \tag{4.4}$$

3.1.6 Sound strength G

In the domain of architectural acoustics, the parameter G, designated as sound strength, serves as a pivotal metric for evaluating the acoustic quality of an enclosed space, with particular relevance to concert halls, theatres, and auditoriums. In accordance with the provisions of the international standard ISO 3382, the term is defined as the differential, expressed in decibels, between the sound pressure level measured at a specified point within the room and that which would be recorded in free field conditions at a reference distance of 10 metres from the same omnidirectional source.

$$G = L_p - L_{p,0} \quad (4.5)$$

In the equation 4.5 L_p is the sound pressure level measured in the room and $L_{p,0}$ is the sound pressure level in the free field, 10 m distant from the reference source.[8]
From a conceptual standpoint, *G* thus quantifies the effect that the environment exerts on the perception of sound strength. Values exceeding 0 dB indicate that the room amplifies the sound in comparison to the free field, while negative values correspond to attenuation.
The interpretation of this parameter is of particular importance in the acoustic design of spaces intended for music and speech. A concert hall of a commendable calibre, for instance, is likely to possess a G-value range of between +3 and +6 dB, which is ordinarily perceived by the audience as a resonant, immersive, and adequately potent auditory experience, devoid of any artificial amplification.

3.2 Materials for Acoustic Applications

The acoustic treatment of a room is designed to control both the direct sound and its reflections from the walls, ceiling, and floor. When anything is immersed in a sound field it will be set into vibration and when sound encounters a surface, it can be transmitted, absorbed, or reflected, depending on the properties of that surface. In order to design a recording studio with optimal acoustics, it is not sufficient to consider only the room's volume and shape; it is also essential to select and apply the appropriate materials so that surfaces can diffuse, reflect, or absorb sound waves according to the desired acoustic response. A conscious and effective use of acoustic materials requires a solid understanding of the different types of materials available and their fundamental behaviors.
Both absorbers and diffusers can be used to prevent acoustic distortion.

3.2.1 Sound Absorption Materials

Sound absorption can be achieved through various physical mechanisms, such as the tortuosity of porous materials, friction between particles, and internal viscous losses. Porous materials are particularly effective at points of minimum acoustic pressure or at high frequencies, when sound has maximum particle velocity. For this reason, recording studios typically use materials such as mineral wool, fibreglass, open-cell foam (polyurethane or melamine) and bass traps, which are usually placed in corners or other critical areas. These elements control reverberations, reflections and standing waves, helping to achieve a balanced acoustic response in the room.
Sound insulation has a different objective: rather than "dampening" the sound inside the room, it aims to prevent sound from being transmitted to surrounding environments. Since absorption-based insulation would be ineffective and require excessive thickness, especially at low frequencies, insulation is instead achieved through mass, layering and structural decoupling. Dense, rigid materials are combined with elastic layers and air gaps to create multi-level systems that can reflect and attenuate sound energy.[16]
In the design of recording studios, this translates into false walls on metal frames containing sound-absorbing material, double or triple layers of high-density panels, suspended ceilings with anti-vibration

elements, and above all, floating floors that decouple the room from the supporting structure. Often, a real 'room within a room' (box-in-box) is created where the floor, walls and ceiling form an independent, dampened envelope that is separate from the building. This drastically reduces acoustic bridges and vibrations transmitted through solid structures.
The amount of insulation required depends on the planned activity and the sensitivity of the intended recordings. Low frequencies are the most challenging, as they propagate easily through walls and floors. Although the human ear is less sensitive to low sound levels in these frequency ranges, this does not negate the need for specific construction systems. It is only through a balanced combination of absorption, mass, and decoupling that it is possible to ensure controlled recording conditions and adequate isolation between rooms.

3.2.2 Sound Diffusion Materials

The use of acoustic diffusers has increased consistently since the 1980s. Historically, convex or otherwise curved architectural elements have been incorporated into the design of concert halls to redistribute incident sound across a wider area and disrupt specular reflections, which would otherwise concentrate acoustic energy in specific zones. As with absorptive materials, the physical dimensions of diffusing surfaces must correspond to the intended operating frequency range. For example, a traditionally designed low-frequency diffuser usually needs a depth of around one quarter of the wavelength to work well.
While diffusion is related to scattering, the two phenomena differ in their characteristics. Scattering generally produces a more irregular redistribution of sound and often depends on the angle at which sound waves strike a non-uniform surface. In contrast, effective diffusion seeks to disperse sound more evenly across space and frequency, producing a uniform acoustic field rather than a directionally biased or randomly scattered pattern.[16]

3.3 Measurement Techniques in Acoustics

3.3.1 Impulse-Response Methods (MLS, Exponential Sine Sweep)

The characterization of a musical environment acoustically is typically undertaken by means of schematization, employing the conventional paradigm of a black box. The black box is the object under examination, the behavior of which is the focus of the study. The device exhibits both input and output characteristics; however, the precise mechanisms underlying its operation remain to be elucidated. In accordance with the principles of acoustic theory, the hypothesis that closed environments are linear time-variant (LTI) systems is postulated. The following text is intended to provide a comprehensive overview of the subject matter.
A system is said to be linear if it satisfies the principle of superposition, which states that the system's response to an input is the sum of the responses to each individual component of the input, with each component applied separately. Additionally, a system is said to be time-variant when the output generated

by a delayed signal is equivalent to the output produced by the original signal, with the delay amount being equal in both cases.
Consequently, to analyze the behavior of an acoustic system that can be schematized as previously outlined, it is imperative to experimentally ascertain the so-called impulse response.
The impulse response is defined as the signal that is generated by a system when a specific impulse signal, mathematically defined as the Dirac delta function, is applied as an input to the system. This specific signal type, when encoded with Pulse Code Modulation (PCM), forms an extensive sequence of zero-amplitude samples, with a single sample positioned centrally that has a value equivalent to the full scale, which is typically taken to be 1.
However, the impulse response measured with a Dirac delta signal can present certain issues, including a poor signal-to-noise ratio, the risk of excitation of nonlinear phenomena due to the use of a signal with a very high slew rate, and the potential presence of very high and very low frequencies in the test signal that can damage the transducers.
Consequently, indirect techniques have been developed to measure the impulse response based on test signals other than the Dirac delta.
Among the indirect techniques, the Maximum Length Sequence (MLS), based on the pseudo-channel white noise signal, was established first, followed by the Linear Sinusoidal Sweep (LSS), which has recently been replaced by the Exponential Sinusoidal Sweep (ESS). The following text is intended to provide a comprehensive overview of the relevant subject matter.

3.3.1.1 Maximum Length Sequence (MLS)

The maximum length sequence (MLS) is a signal employed to measure two-port systems. When reproduced by a loudspeaker, the auditory effect is that of white noise: a signal that appears random and non-periodic, requiring a sufficiently long observation time to accurately estimate its spectrum.
In essence, MLS falls within the category of pseudo-random signals, characterized by the repetitive cyclical recurrence of a predetermined sequence of defined duration. These signals possess characteristics that are strikingly similar to random noise, to the extent that they are frequently indistinguishable.
An MLS signal can be conceptualized as a sequence of positive and negative pulses of equal amplitude, distributed in a pseudo-random manner. In comparison to the conventional test signal, which is based on a solitary pulse per measurement interval, MLS comprises a greater number of pulses and, consequently, possesses a higher overall energy content.
The measurement principle is straightforward: the MLS signal is applied to the system input and correlated with the output signal to obtain the impulse response. To ensure the reliability of the result, the length of the sequence must be a minimum of equal to the reverberation time of the environment under analysis. Subsequent to the acquisition of the impulse response, the frequency response of the system can be readily obtained through the implementation of a Fourier transform (FFT).
Following an extensive period of research focused on the efficacy of the MLS approach, it has been observed that distortions emerge in the impulse responses derived from this methodology, thereby influencing the accuracy of reverberation time measurements.

The Periodic Impulse Response reveals irregularities in portions of the signal where a homogeneous decay would normally be expected. Similarly, in the Schroeder curve, the ideal exponential decay is disrupted by broadband distortion as well as by low-frequency irregularities.
Additional sources of nonlinearity, such as transversal distortion, variations in the initial gain, clipping phenomena, and limitations imposed by the slew rate, further compromise the reliability of MLS-based measurements. These factors have significantly reduced the adoption of MLS techniques in contemporary professional acoustic practice, where alternative measurement methods are generally preferred.

3.3.1.2 Exponential Sinusoidal Sweep (ESS)

This technique utilizes a sinusoidal sweep as the test signal, whose frequency undergoes exponential increase over time. A foundational element of this approach is that the deconvolution process is no longer circular, as is the case with MLS, but is instead executed directly in the time domain.
The action of the system on the test signal can be expressed as a covolution of a delta of Dirac with the filter. Therefore the impulse response will be the charachteristic of the filter.
The inverse filter can be determined with ease by employing the same test signal that was reproduced in reverse. However, if the sweep is non-linear, its spectral content will not be flat. In such cases, corrective equalization must be applied to compensate for spectral variations and obtain a frequency response of the impulse response that is as regular as possible. Using a non-linear or logarithmic sine sweep with the respect to a linear one has two main advantages: a better energy distribution on the spectrum, since the logarithmic sine sweep holds more energy on the low frequencies compared to the high frequencies, considering that the progress is slower on the lows and faster on the high ones. In the end the resulting spectrum pink instead of being white. Having a warping linear system it is still possible to study the warp tendencies When a system introduces distortion, the input of a signal generates harmonics, i.e., components at frequencies that are multiples of the test signal frequency. Consequently, this harmonic distortion manifests itself in responses associated with these multiple frequencies.
To obtain the impulse response, these components are represented as straight lines that are "rotated" on the plane until they are reduced to a Dirac impulse. The straight lines corresponding to distorted behavior undergo a greater rotation than the linear ones. For this reason, the distorted impulse responses appear temporally before the linear IR. The distinction between distorted and linear terms is evident due to the logarithmic nature of the frequency sweep, which ensures constant time intervals between each frequency and its first harmonic. This property enables the analysis of individual harmonic distortions separately, their summation, and the subsequent determination of the total harmonic distortion. In conclusion, by isolating solely the response associated with the linear components on the right, the distortion-free impulse response is derived.
In our objective tests, we chose to adopt the latter technique, as it provides a significantly improved signal-to-noise ratio and greater immunity to system distortion. In conclusion, the Exponential Sine Sweep (ESS) method offers the important advantage of enabling system analysis without assuming linearity and time invariance (LTI). This makes it applicable even in cases where the system exhibits pronounced nonlinear behavior or time-variant characteristics.[17]

3.4 Standards and Best Practices

3.4.1 Stage acoustic (ISO 3382-1)

The measurement of stage acoustics is typically undertaken in the context of orchestra stages and large halls, with the objective of evaluating the acoustic condition experienced by musicians during their performance. In reality, the occurrence of such a condition is improbable, given the number of musicians and the orchestra that can be accommodated within the designated recording space. Nevertheless, we sought to consider the acoustic perception of the musician within our ambient environment. [2]

3.4.2 STI model IEC 60268 – 16:2020

The Speech Transmission Index (STI) model is an objective method for assessing the comprehensibility of a sentence or vocal communication when transmitted through a communication channel. In this case study, the technology is employed to predict the intelligibility of speech within a recording studio, which constitutes the medium of vocal transmission that forms the focal point of this thesis. The acoustic characteristics of this medium will be the primary focus of our study.

The STI model is predicated on the premise that the speaker employs a standardised vocal timbre, articulates clear speech, and articulates at a rate of 3-4 syllables per second, in the presence of a listener with normal hearing.

It is important to note that when a person speaks, the sound intensity is not constant, and these fluctuations in volume are important for correct intelligibility.

Variations in speech intensity are termed 'modulations' and can be described as a function of modulation frequency, giving rise to the modulation spectrum. In well-articulated speech, modulation frequencies typically range from 0.5 Hz to approximately 16 Hz, with a peak around 3 Hz, which corresponds to the normal cadence of human speech. By comparing the signal envelope spectra recorded directly from the speaker's voice with those detected after transmission through a certain channel, it is possible to assess how much the original fluctuations have been attenuated. Any alteration of the modulation spectrum introduced by the channel is interpreted as a reduction in modulation depth at one or more modulation frequencies. This degradation is expressed by a modulation transmission value calculated for each octave band that constitutes the speech spectrum. By appropriately choosing the shape of the test signal, it is possible to obtain an effective signal-to-noise ratio that considers not only background noise, but also temporal distortions and any non-linearities of the system.

The standard test signal comprises a carrier with a spectrum modelled on the human voice, to which a sinusoidal modulation with a frequency of f_m is applied. In the comprehensive STI (Speech Transmission Index) methodology, the utilisation of noise signals across seven octave bands, ranging from 125 Hz to 8 kHz, is employed. The analysis involves a total of 98 measurements, resulting from the combination of the 14 modulation frequencies with the seven frequency bands, and the entire procedure can take up to 15 minutes. The STIPA method has been introduced to simplify the process. This method uses a single test signal containing a predefined combination of two modulations for each octave band. This accelerated approach is especially well-suited for the evaluation of speech transmission in real-world environments and in sound reinforcement systems and in fact is the one used in this thesis.

4 History and background of the case study

The case study presented in this thesis focuses on the recording room of the production studio located within the SERMIG Arsenale della Pace in Turin, Italy. This distinctive locale unites a profound historical legacy with a contemporary mission of peace and cultural engagement.
Originally constructed in 1580 as a gunpowder factory, the edifice underwent a series of transformations over the centuries, ultimately evolving into the Arsenal of Artillery of Turin in 1852, a development precipitated by a catastrophic fire.
In 1983, the derelict arsenal was transformed by a group of young volunteers led by the SERMIG (Servizio Missionario Giovani) community into what is now known as the Arsenal of Peace. The objective of this initiative was to transform this emblem of warfare into a domain conducive to peace, solidarity, and social action. In the contemporary era, the Arsenale functions as a focal point for intercultural discourse, educational endeavours, spiritual practice, and the provision of humanitarian assistance. The facility operates on a 24/7 basis and provides assistance to individuals facing various challenges, including single mothers, former prisoners, and migrants. The services offered include shelter, medical support, and vocational opportunities.

4.1 First Acoustic Enhancements

The establishment of the SERMIG Production Studio within this space was with the objective of providing support for the community's musical and multimedia initiatives. Since 1998, the organisation has also welcomed external clients, thereby achieving a balance between social purpose and the provision of professional-grade production services. The studio has a control room, a recording room (the subject of this thesis), designed originally in 2000 by Ing. Brosio, suitable for both ensemble recording and live performances. The room configuration has been improved by the studies and projects of Doct. Fringuellino, with a control room Dolby Atmos with a 7.2.4 speakers set up and with an acoustically optimized recording room.
The studio's mission is deeply rooted in the performance and recording of acoustic music, often involving live musicians, choirs, and ensembles. In circumstances where it is possible, productions are encouraged to opt for traditional recording methods over the use of virtual instruments. It is evident that, over time, the studio has produced soundtracks, television content, educational material, and collaborated with national broadcasters and famous artists.

4.2 Current Acoustic Condition of the case study

The room, devoid of any furnishings, exhibits harmonious proportions. Its height is recorded as 3.00 meters, with a length that varies slightly. At its widest point, the object measures 7.32 meters, narrowing to 7.10 meters at its narrowest point. The width also exhibits this slight irregularity, ranging from a

maximum of 5.05 metres to a minimum of 4.86 metres. The total perimeter of the space is 24 metres, and its area is approximately 35 square metres. The total volume of the space is thus 105 cubic metres.

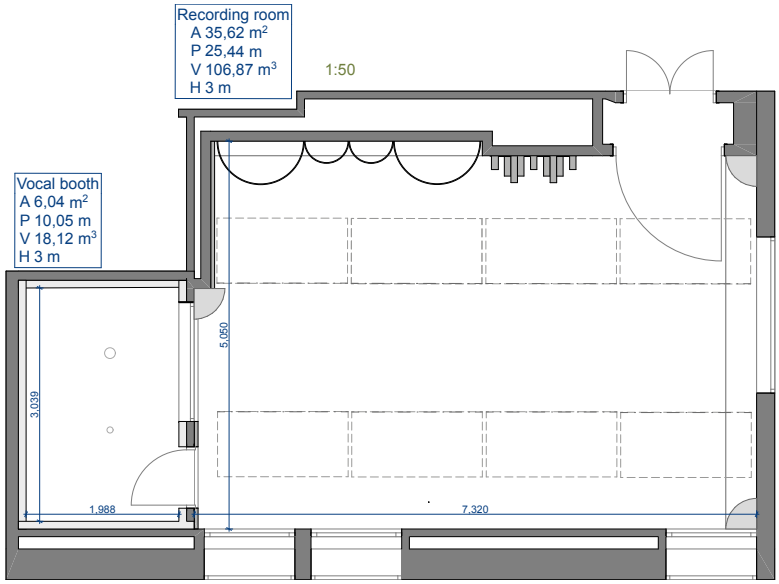


Figure 4.1 SERMIG's live room floorplan

The room contains several elements that have a significant impact on the internal acoustics, including three windows, two metal doors, one larger for the main entrance and one smaller leading to the booth, a glass panel for the director's view and another for the booth's view. The floor is composed of parquet, and along the entire wall of the director's view there is a wooden bench that contributes to the distribution and reflection of sound.

The ceiling is composed entirely of perforated sound-absorbing plasterboard, likely of the Knauf Danoline Plaza type. A large cavity, with a thickness varying between 30 and 35 cm, is present behind the perforated surface. This cavity is partially filled with rock wool. Above this system are 9 mm-thick wooden sails, inclined in a "V"-shape with respect to the floor and arranged along the long side of the room. The modules are located at a height of 286 cm from the floor and arranged in two rows of four modules each, for a total of eight panels, each measuring 170 x 90 cm. In the most favourable conditions for diffusion, characterised by a significant separation between the receiver and the sound source, the sails initiate the process of sound diffusion, commencing at approximately 280 Hz and gradually descending towards lower frequencies. The upper, mirrored part of the structure has been shown to enhance the lateral efficiency of acoustic diffusion. Additionally, the rear-facing design of these devices obviates the membrane effect, thereby precluding the undesirable absorption of low frequencies.

In the corners of each compartment, conical-section pressed polystyrene elements have been meticulously placed at the point of intersection between the walls and the ceiling.

Opposite the windows, there are four large cylindrical diffusers on the wall. These diffusers have a curved section that is designed to promote sound diffusion. The installation of these units is performed on a superstructure positioned atop an exposed brick base, a configuration that serves to enhance the structural integrity of the entire assembly and ensure stability. The internal configuration of the elements is such that

they are hollow, enabling air circulation at the rear. This reduces the membrane absorption effect, which has the potential to interfere with low frequencies. Constructed on the opposing wall to the one adjoining the control room, this stone wall was erected with the objective of enhancing sound diffusion and achieving a more balanced acoustic environment.
A Schroeder diffuser, which is identical to those located in the rear corners of the control room, is installed on the wall between the access door and the cylindrical diffusers. This choice facilitates an aesthetic connection to the other room while simultaneously enhancing sound diffusion within the hall.
The diffuser is 215 cm in height and positioned approximately 10 cm from the floor, so that its uppermost extremity is aligned with the superior edge of the access door.
The object under discussion is a finely finished, acoustically reflective wooden element designed to diffuse sound evenly. In this case, the diffusion is based on a sequence of quadratic residues, which allows for a homogeneous cylindrical diffusion of the incident sound waves within a certain frequency range and a uniform distribution of sound energy in a plane orthogonal to the vertical axis of the diffuser. Schroeder identified the use of quadratic residues as the ideal solution to ensure this directional uniformity.

The QRD diffuser panel, in Figure 4.2, has been engineered to ensure optimal efficacy within the frequency range spanning from a lower limit of 285 Hz and an upper limit is 2000 Hz. The component in question has a fairly compact overall footprint, with a wood thickness of 1 cm, a width of 110 cm, a depth of 35 cm, and a height of 210 cm. The 8.5 cm wide slots contribute to effective sound diffusion, while the interior is filled with porous sound-absorbing panels to improve overall acoustic characteristics.
Aesthetically speaking, the structure is characterised by a series of adjacent grooves of varying depths, which have been carved out of a vertical reference plane. In the configuration under consideration, the diffuser consists of three sequences (termed "periods"), each formed by N = 7 grooves of uniform width w = 8.5 cm: it follows that each period is 55 cm wide.
The seven grooves in each period exhibit varying depths, determined by the sequence of quadratic residues of the prime number N = 7. This sequence governs the geometry of the panel and defines its diffusion characteristics.

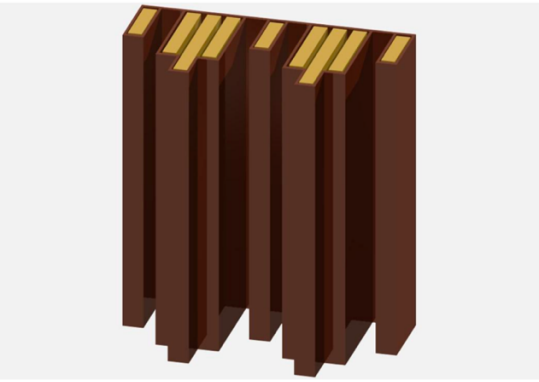


Figure 4.2 The rendering illustrates the Schroeder diffuser as it would appear under examination, with the internal filling clearly delineated in the sectioned top.

5 Objective measurements

5.1 Measurement in the Anechoic Chamber

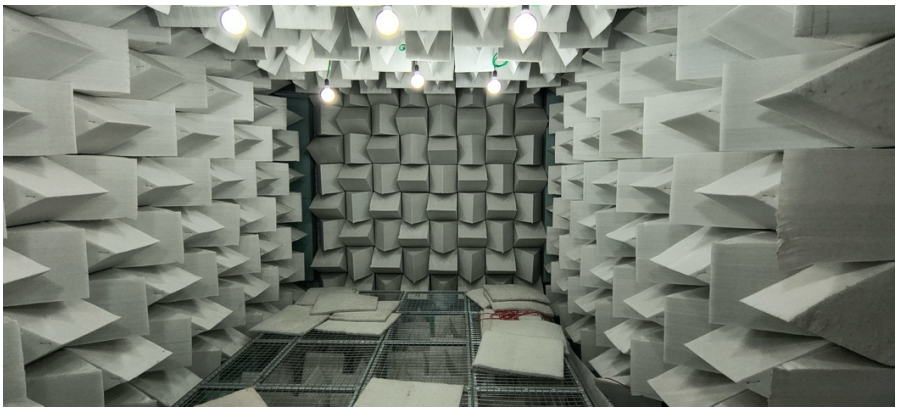


Figure 5.1 Anechoic chamber of DENERG department of reasearch in Politecnico of Turin

5.1.1 Equipment and Setup

The equipment used for acoustic measurements included the following devices LARSON DAVIS L&D824, integrating sound level meter, serial number 3641, Class 1 which includes a L&D2541 ½-inch free field condenser microphone capsule and L&D PRM902 preamplifier, in accordance with the following regulating forces:

- Procedures D0001.8046.
- IEC 61672-1:2007 Class 1;
- IEC 60651-2001, IEC 60804-2000, ANSI S1.4-1983 Type 1 for 1/3. 1/1 Octave filters;
- ANSI S1.11-1986 Type 1C;
- The IEC 61260-am1-2001 standard is classified as Class 1.

The condenser microphone has a frequency response shown in Figure 5.2.

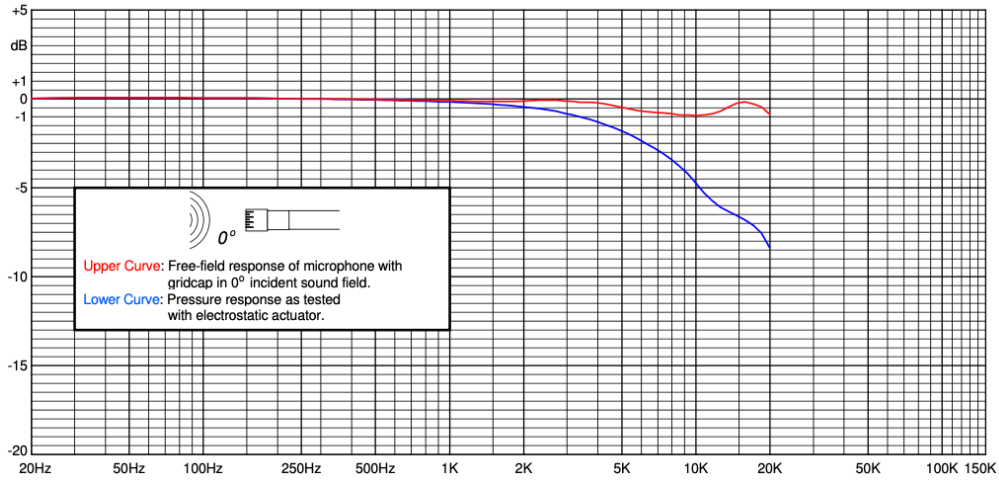


Figure 5.2 L&D2541 ½-inch free field condenser microphone capsule's frequency response

Together we used a LARSON DAVIS CAL200 calibrator, serial number 5235, complies with current Class 1 regulations:

- Procedures D0001.8190.
- The IEC 60942:2004 standard classifies devices according to their water resistance, assigning them Class 1.

The LARSON DAVIS L&D824 and the calibrator LARSON DAVIS CAL200 instrument was officially calibrated on 3^(rd) July 2025.

For sound emittion it has been used a NOISEMETER TWELVE NMTW1 dodecahedral spherical source, serial number TW000105, complies with standards ISO 140 and ISO 3382, with a NOISEMETER WEP BOOSTER power amplifier designed for the spherical source.

While for the low frequencies we used a PHON-X SB-1 passive subwoofer, with serial number PXSB100075. To bypass all the equipment frequency responces we used a KLARK TEKNIK DN 360 multiband stereo equaliser.

All the process of measurement has been recorded and carried on with an HP Pavilion DV7 laptop, serial number HPMH-648102-061, equipped with AURORA software for impulse response measurement and analysis and a FOCUSRITE SCARLETT 8i6 external sound card, serial number UF6207412336. All the equipment explained is shown in Figure 5.3.

To ensure optimal routing, it is imperative to verify that the personal computer (PC) is connected to the external audio card, which is designated as input channels A and B of the equaliser. The amplifier's output is then routed as inputs to the equaliser. It is evident that the amplifier outputs will have the channel B supplied for the subwoofer's alimentation, while the channel A will be directed to the dodecahedral source. It is evident that the latter will be required to supply a weaker power than that required by channel B of the amplifier. This is due to the fact that the spherical source has already integrated its amplifier.

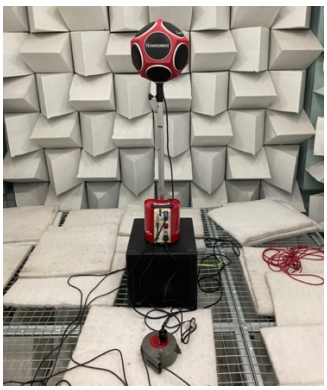
In order to take measurements of the sound level, it is necessary to position the sound level meter at a distance of 10 metres from the sound source. In instances where this configuration is not feasible, the distance must be recalibrated. In the case of this thesis, the distance is 2.5 metres.

The sound level meter can be connected to a personal computer in order to facilitate real-time corrections.

The calibrator must be set at 114 dB and supply a 1kHz tone to the sound level meter.

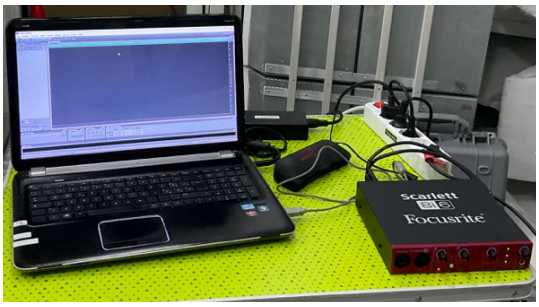
In the event that the calibrator and sound level meter are unable to maintain the reference level of 114.5 dB, it is necessary to calibrate them. In order to ensure that the instruments maintain level accuracy within a maximum error margin of ±0.5 dB between the start and end of the measurement, it is imperative that this operation is performed at two-year intervals.

The random connection of the sound level meter must be activated solely in cases where measurements are being taken in spaces with reflections; this is not necessary in an anechoic chamber. This correction is advantageous in that it addresses the discrepancy between open-field microphones, which are typically utilised in American manufacturing, and non-open fields, which are more commonly favoured in European manufacturing, as per ISO standards. The sound level meter and its microphone, being of American manufacture, are classified as open-field microphones according to ASTM standards.



(a)

(b)



(c)

(d)

Figure 5.3 (a) NOISEMETER TWELVE NMTW1 dodecahedral spherical source, (b) KLARK TEKNIK DN 360 multiband stereo equaliser, (c) HP Pavilion DV7 laptop with the FOCUSRITE SCARLETT 8i6 external sound card, (d) LARSON DAVIS L&D824 and the calibrator LARSON DAVIS CAL200

5.1.2 Results of the calibration

Following the establishment of a connection between the calibrator and the sound level meter, it is imperative to ascertain that the audio levels of the two sources are elevated to the maximum extent possible in order to attain the maximum sound pressure level. This is to ensure that the reproduction of a pink noise sample does not result in clipping.

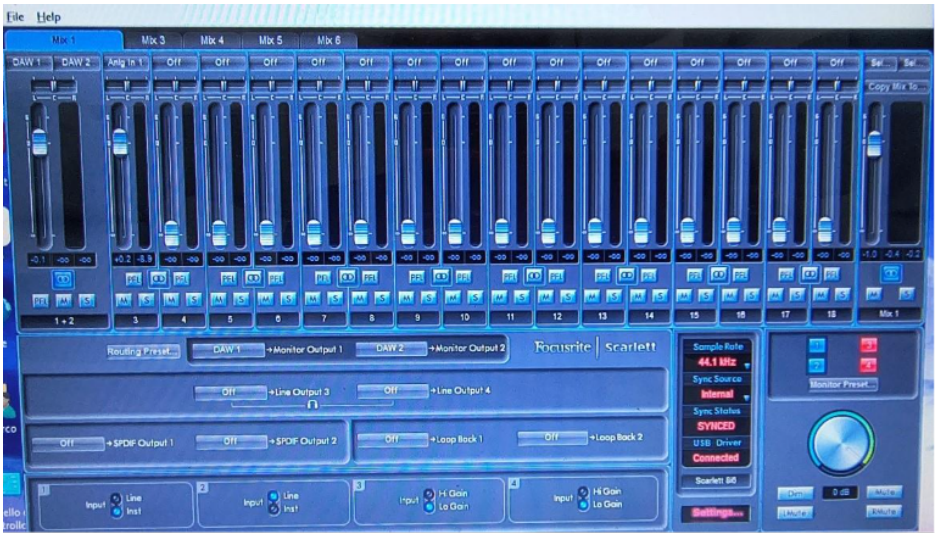
The setting of the gain of the phonometer and of the external audio card from the pc display must be at 0 dB, with the Scarlett main monitor level at 6.5, the acquisition volume of the microphone 1 on the Scarlett at 6.5 and the pc volume bar to the Scarlet at 96 over 100.

Then the knob of the channel A oriented at 3 o'clock and the channel B to the 10 o’clock, while the WEP booster of the dodecahedrical at 3 o’clock.

Concearning the Equalizator, the start setting has to be as follow: channel A with the filter on the 30 Hz and on 6 dB scale, while the channel B with no filter an set on 12 dB scale.

While making correction on equalization we should pay attention that each third of an octave philter, having a bell shape, has influences the immediately adjacent filters.

(a)



(b)

(c)

(d)

Figure 5.4 (a) sperical source's NOISEMETER WEP BOOSTER power amplifier settings, (b) settings from the DAW pc display, (c) settings of the gain knobs of the audio card, (d) final setting of the equalization to obtain a flat response

Subsequent to the measurement of pertinent parameters with the sound level meter, a graph was produced which demonstrated the trend over time of the equivalent average sound pressure level (Leq). The application of this graph facilitated the implementation of appropriate corrections to the acquired data, thereby compensating for any discrepancies that may have arisen due to instrumentation or environmental factors. Subsequent to the implementation of these corrections, it was possible to ascertain an overall equivalent level of $85dB_{SPL}$.

The Figure 5.5 shows the plot of the frequencies over the equivalent level of sound pressure with a flat equalization. In other words, the graph shows the dodecahedral source and its subwoofer frequency behaviour without calibration.

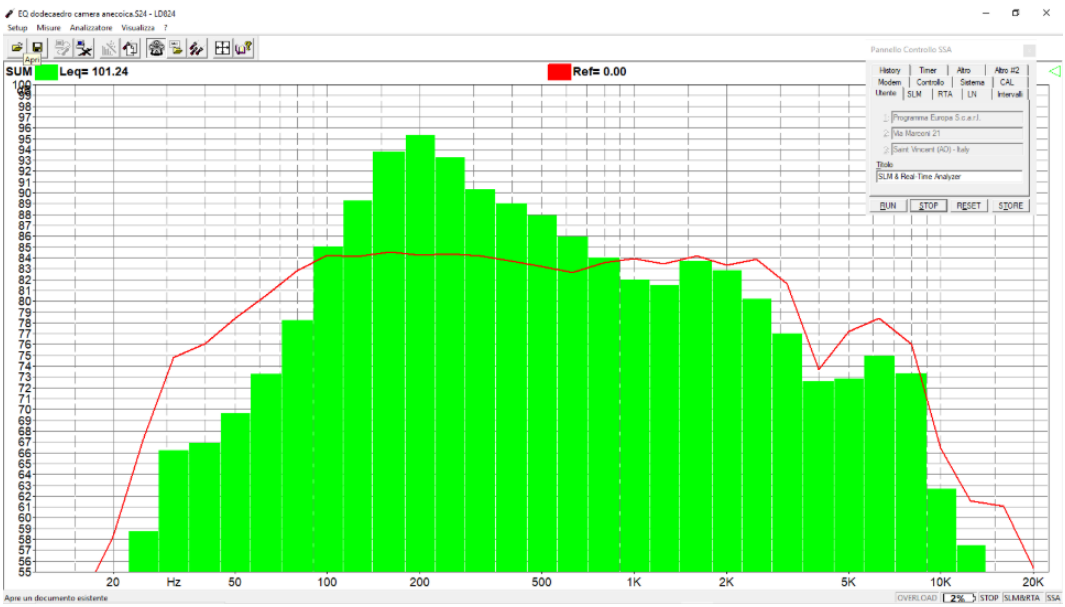


Figure 5.5 the plot of the frequencies over the equivalent level of sound pressure with a flat equalization

While Figure 5.6 shows the sources calibrated with the equalization working with setting in Figure 5.4 (d).

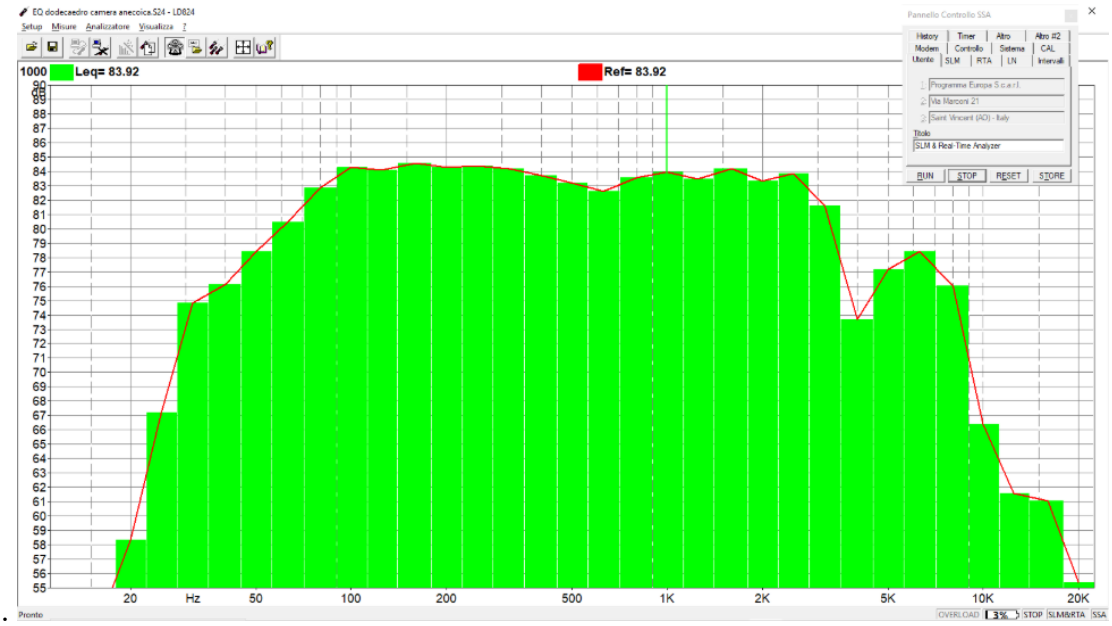


Figure 5.6 sources calibrated

5.1.3 Impulse Response

We proceeded then with the measurements of the impulse response, using the exponential sine sweep technique, for which the sound level meter must be connected mono to the external audio card. Convolution of recordings made with identical parameters to those employed in the creation of the signal to be convolved, and subsequently transmitted to the room, is the only viable option. To illustrate this principle, consider a sine sweep generated at 44 kHz and 16 bits. To ensure accurate convolution, it is essential that an audio file of identical parameters is recorded in the room, i.e. at 44 kHz and 16 bits. Initially, the DC/AC output of the LD824 sound level meter must be connected to the input of the Scarlett sound card. In the event that the Scarlett is not available, it is possible to use the microphone input of the PC's sound card as an alternative. The subsequent stage of the process is to connect the Scarlett's monitor output, or, in the event that this is not available, the computer's headphone output, to the amplification system that will broadcast the sound into the environment. Subsequently, it is imperative to ensure that the microphone input is correctly enabled using the sound card mixer.

In regard to the Scarlett routing, DAW 1 goes with Output 1 and DAW 2 goes with Output 2, without utilising the MIX in order to avoid the occurrence of signal feedback.

In Adobe Audition, it is necessary to create a mono sine sweep file, whilst ensuring that the system is not connected to the power supply in order to avoid any potential damage. The sweep file should then be inserted into the first track, and the second track should be prepared for mono recording. Subsequent to the recording of the signal, the convolution with the inverse filter of the sweep, which should already have been saved to the clipboard, should be performed.

The signal utilised encompassed the entire audible spectrum, ranging from 20 Hz to 20 kHz, with a duration of 10 seconds, repeated thrice at 5-second intervals.

The audio recordings were captured at a sampling rate of 48 kHz, employing 32-bit floating-point precision. In conclusion, the sine sweep can be viewed in Adobe Audition by utilising the time-frequency mode, otherwise referred to as 'spectral frequency'. This enables the user to discern the frequency produced at any given moment. Through the analysis of the recorded signal, it is possible to identify any resonances within the environment.

However, in order to achieve a higher degree of accuracy, it is recommended that the theoretical formulae which describe the relationship between frequency and time in the sweep signal be consulted, look at equation 6.1.

As expected, the impulse response from an anechoic chamber obtained is nearly the direct sound pulse followed by a period of silence, as the chamber's design absorbs nearly all sound, preventing reflections from reaching the microphone.

$$x(t) = \sin\left(\frac{\pi f_{inf} T}{\ln\left(\frac{f_{sup}}{f_{inf}}\right)}\left[exp\left(\frac{t}{T}\ln\left(\frac{f_{sup}}{f_{inf}}\right)\right) - 1\right]\right) \tag{6.1}$$

The equation shown in 6.1 is the exponential sweep, where f_{inf} is the starting frequency (usually of 20 Hz), f_{sup} is the ending one (usually 20k Hz) and T is time-period of the sweep (tipically $10 \div 50\ s$). It is evident that the sine-sweeps are more efficient for measuring the transfer function in acoustics for a variety of reasons. Primarily, the response to a single sweep that is not repeated is almost equivalent to that which would be observed with periodic repetition, thus resulting in a shorter measurement period.

5.1.4 Definition of G

In circumstances where the measurement of G in a reverberation chamber according to ISO 3743 is not feasible, or where an intensity measurement system compliant with ISO 9614-2 cannot be employed, an alternative procedure is to be adopted. This alternative procedure is based on measurements taken in an anechoic chamber.

In order to ascertain the gain of the room, a reliable reference signal is required. As specified in ISO 3382-2 (2006), Performance Spaces, this reference must be represented by an impulse response which is to be acquired in an anechoic chamber situated at a distance of 10 metres from a spherical source which has been previously equalised.

In practice, the Aurora software is utilised to equalise the spherical source in the anechoic chamber prior to the measurement of a normal impulse response. In order to proceed with the subsequent stages of the process, it is necessary to open the "Acoustical parameters" window (see Figure 5.7). This should be followed by selecting the "Store G reference signal" function (see Figure 5.8). These actions can be found under the Aurora effects menu.

The graph screen is substituted by a dialog box in which the measurement distance can be set. The actual value should then be entered into the field, with the value in this case being 2.5 m. The "OK – Store" function should then be selected to confirm.

The G values calculated by the "Acoustical parameters" plug-in have been observed to converge towards the actual values; however, this is only guaranteed under the condition that the sound source is point-like and perfectly omnidirectional. It is therefore important to emphasise that the measurement of G is particularly delicate and that obtaining a rigorous estimate from the impulse response is often complex.

The anechoic reference measurement constitutes the most critical phase of the process. It has been frequently observed that a slight rotation of the "omnidirectional" dodecahedron, even of a few degrees, can produce variations in the amplitude of the free-field signal of the order of a few dB, especially above 1 kHz. This highlights the imperfect angular uniformity of the source.

Finally, it is imperative to acknowledge that throughout the entire IR acquisition process, there is a significant risk of compromising the absolute calibration of the sound pressure level (SPL) by introducing uncontrolled gains or automatic scaling. It is imperative to meticulously calibrate the G value, a process that demands precise control over recording levels and the application of scaling factors during deconvolution. Any deviation from this protocol must be avoided to ensure the accuracy and reliability of the results.

In order to facilitate the correct export of data, it is recommended that a series of preliminary procedures be followed.

Firstly, it is necessary to ensure that the decimal separator is set to a period.

This configuration is pre-established in Aurora; however, in the Windows operating system, it necessitates manual adjustment through the regional settings.
Secondly, when exporting data from the Audition clipboard to Aurora, it is recommended that an Excel file be opened and that all cells be formatted as text. This is an essential step in preventing misinterpretation of values, particularly those that include the decimal separator, and thus ensures correct data management and consistency.
The value of G was determined by requiring that, at a distance of 10 metres from the sound source, the gain detected by the microphone be equal to 0 dB. In this instance, subsequent to scaling according to the actual distances, a value of G equal to 12 dB at the considered distance was obtained.
This value is now employed as a calibrated reference for the subsequent evaluation of the gain inside the SERMIG recording studio.

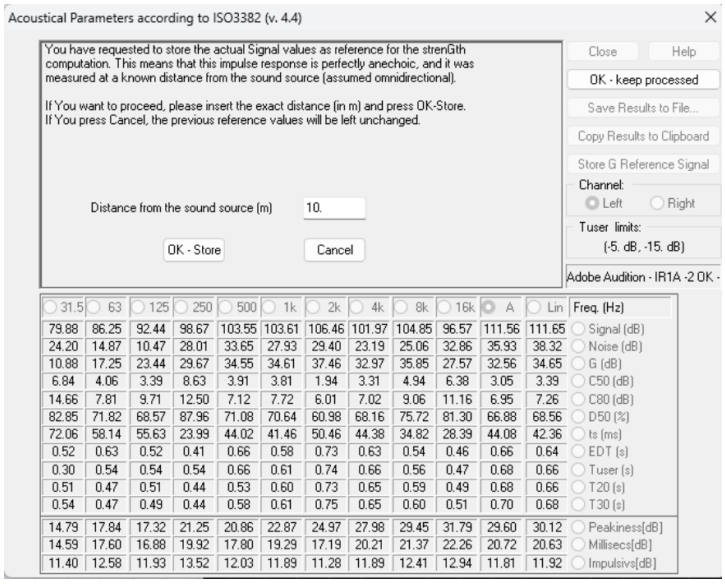


Figure 5.7 acoustical parameters window

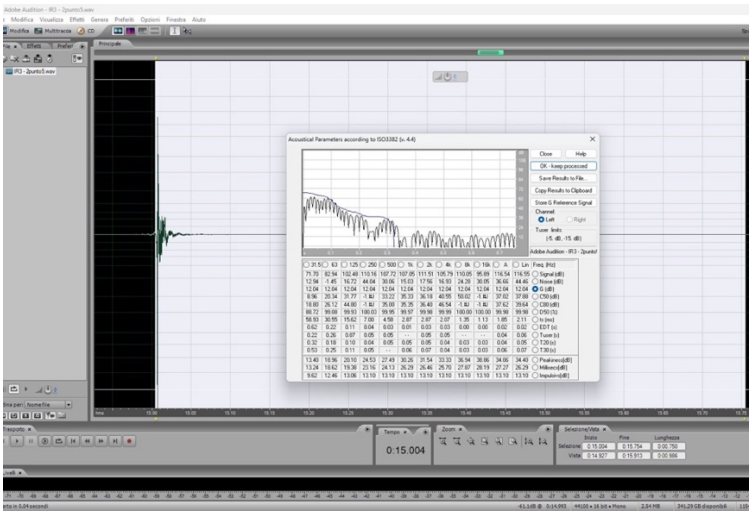


Figure 5.8 storing of the G reference

5.2 Main Room Measurements

5.2.1 Equipment and setup

The measurements were obtained using the same configuration and apparatus employed in the anechoic chamber tests, ensuring equalization and gain values remained constant to yield results independent of the equipment. Subsequently, additional immersive recordings were made using an Insta360 camera and an O-Zylia microphone, both connected directly to a MacBook Air M3 (serial number K1Q657PF29) and controlled and recorded using the desktop application provided by the manufacturer.

In additional, two professional microphones were utilised for the stage acoustic measurements. The first is a class 1 measurement microphone, model NTi Audio M221, which offers a very wide frequency response, from 5 Hz to 20,000 Hz. The device is an omnidirectional microphone, characterised by its compact size (150 mm in length and 20.5 mm in diameter) and its ability to capture sound from all directions. The device is capable of withstanding sound pressure levels of up to 144 dB, possesses a balanced impedance of 100 Ohms, and necessitates 48 V phantom power. The apparatus is also distinguished by the presence of a removable capsule and a sturdy metal diaphragm. These characteristics are indicative of the instrument's reliability and accuracy.
In this study, the authors opted for a more economical microphone, the Behringer ECM8000. This electret condenser microphone is characterised by its precision and extremely linear frequency response, which covers the range from 15 Hz to 20 kHz. In addition, the device under consideration functions as an omnidirectional balanced receiver, a configuration that renders it well-suited for the implementation of room correction applications. This is due to the fact that it captures sound in a uniform manner from all directions. However, it is not designed to handle extreme sound pressure levels like professional high-SPL models, and since it is not Class 1 certified, it does not have an individual calibration curve. The device exhibits a balanced impedance of 600 ohms and necessitates 48V phantom power.
The prevailing environmental conditions underwent minor fluctuations during the specified period. The temperature exhibited an increase from 24.2°C at the initial stage to 24.4°C at the concluding stage, while the humidity level experienced a decline from 67.7% at the beginning to 63% at the conclusion of the period under consideration.
The configuration of the recording room was such that the acoustics were already established, and several fixed instruments were present which could not be moved. These instruments included a motorised Leslie organ, a Rhodes piano, a grand piano, and acoustic measurement equipment. Prior to the commencement of the survey, the temporary instruments, including the drums, and other superfluous items, were removed. It should be noted that only one person was present at the time of measurement.
The measurment has been taken following the schema in the floor plan in the next paragraph.

5.2.2 Impulse Response of the recording room

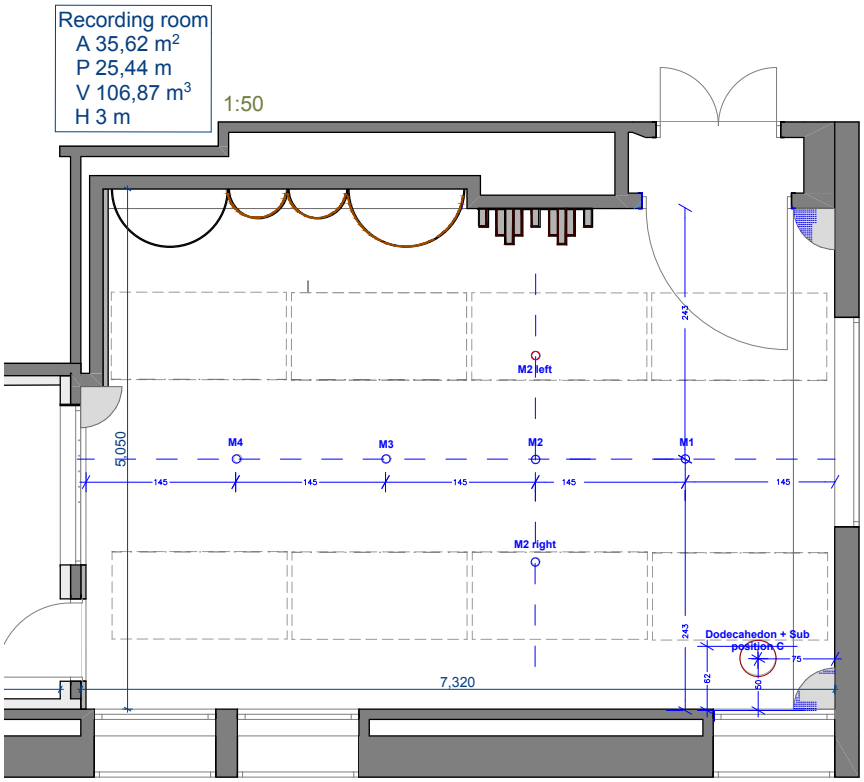


Figure 5.9 floorplan showing the receivers and the source position for the IR measurements

The impulse response of the room where recorded having the dodecahedral spherical source and the subwoofer in what we call Position C, which has been developed for monaural impulses and the Zylia system, is located 75 centimetres from the control room wall along the x-axis (the longitudinal axis) and 50 centimetres from the window wall along the y-axis (the transverse axis). The height of the detection point, on the z-axis, is set at 1.5 metres from the floor.

The measurements obtained using the Larson Davis L&D824 sound level meter were repeated at the same six points that were utilised for the preliminary recordings with the Zylia microphone. These preliminary recordings were analysed in paragraphs 5.4 to 5.4.2. The purpose of this analysis was to acquire monophonic recordings from which the acoustic parameters necessary for characterising the environment could be derived. Consequently, this section will exclusively analyse the results obtained through the utilisation of sound level meters.

The impulse response was measured using an exponential sine sweep as the test signal, which was repeated three times, in order to preserve the signal to noise ratio, at each measurement location. The resulting acoustic response was recorded in the room using Aurora.

5.2.3 Data elaboration

Following the conclusion of the measurement session, a total of six recordings were obtained, with a sampling frequency of 48 kHz. These recordings were then exported in WAV format, with each recording comprising three repetitions of the signal.
The calculations were performed using a MATLAB script. In order to extract the system response from the recorded signal, the recordings were convolved first with the inverse of the sweep used. Consequently, the convolution process effectively "erases" the original sweep, thereby producing the impulse response of the system. This impulse response constitutes the fundamental data for all subsequent analyses. Absent this step, it would be impossible to differentiate the influence of the signal from that of the system itself.
Subsequently, the signal repetitions, generated to improve the signal-to-noise ratio, were isolated into separate tracks and analysed individually.
The acoustic parameters of the room were then calculated for each of the three repetitions relating to the six measurement points. The results were organised by octave band and saved in an Excel file. Using the values obtained for each repetition within the same file, it was also possible to determine the mean and standard deviation of all the acoustic parameters considered for each measurement point. For the acoustic characterisation of the SERMIG recording room, the parameters analysed in this thesis are: EDT, reverberation times (T_{20} and T_{30}), clarity indices (C_{50} and C_{80}), definition parameter (D_{50}) and centre time (T_s).

5.2.4 Experimental results and comparison with literature case studies

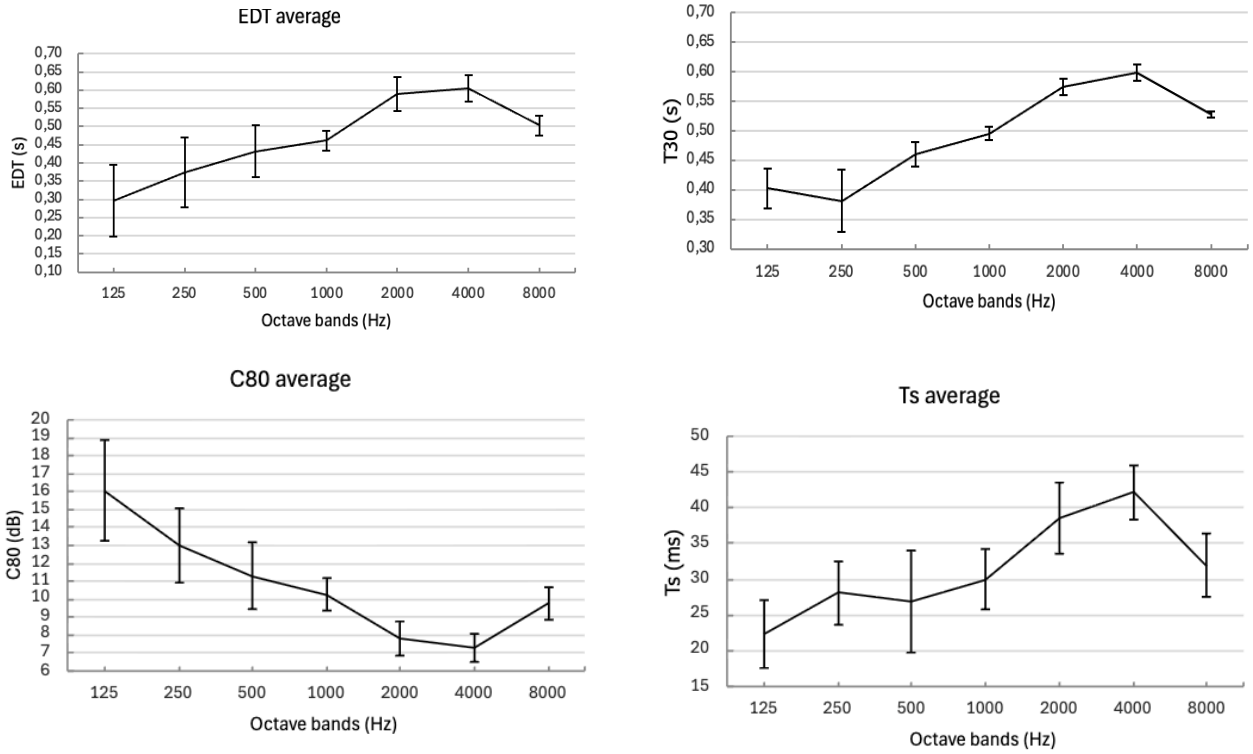


Figure 5.10 average values of key parameters of frequency of the objective measurements

Following a comprehensive review of the extant literature, it is evident that the most significant parameters for describing the acoustic behaviour of a space are EDT, T_{30}, T_s, and C_{80}. These parameters facilitate the evaluation of the initial decay of sound energy, the overall reverberation time, the temporal distribution of energy, and the ratio between early and late energy, respectively.

In the case in question, the T_{30} interval, which remains between 0.4 s and 0.6 s from low to high frequencies, places the room, according to the classification proposed by Lauer-Münchhofen[18], in an intermediate position between the second and third types. It can thus be posited that the space under consideration may be interpreted as a live room, characterised by a slightly more reverberant timbre at high frequencies. In accordance with the findings of Tavelidou's study, this interpretation is substantiated by a comparison. The measured values are positioned between the configurations identified by Tavelidou as B and C. Configuration B is distinguished by a highly reverberant and clear response, albeit with relatively uniform frequency. In contrast, configuration C exhibits more contained reverberation, accompanied by a more pronounced timbre, attributable to the disparity between low and high frequencies.[19]

With regard to EDT, in acoustically balanced environments this parameter tends to be similar to T_{30}, since the energy decay curve maintains an almost exponential trend.[3] In this instance, EDT is marginally higher than T_{30}. As Barron has observed, this condition indicates that a significant portion of the sound energy is concentrated in early reflections.[20] The result is that the subjective perception of reverberation may be longer than suggested by T_{30} alone. [3] The trend of the EDT curve demonstrates a strong similarity to that observed by Tavelidou in configuration B, which represents the most reverberant case among those analysed in her study.

The space also has relatively high C_{80} values. However, when considering the range of values measured in this study and comparing them to the range of values reported in the literature on recording studios of similar volume, as analysed by Lauer-Münchhofen, where values of around 30 dB are also reached, it is evident that the range measured in this study is only slightly higher than the values generally recommended in the literature.[5]

Finally, with regard to Center Time (T_s), several reference studies in architectural acoustics [3,20,21] indicate a typical range of between 20 milliseconds and 80 milliseconds for recording rooms. The values measured in the analysed space fall fully within this range, oscillating between 19 milliseconds and 42 milliseconds. This indicates a temporal distribution of energy that is compatible with good sound definition.

5.3 Stage acoustic

5.3.1 Importance of the stage acoustic

It is reasonable to hypothesise that the rationale behind the decision to undertake acoustic stage measurements within a music recording studio may raise some curiosity. This is, in fact, a valid and methodologically sound consideration. In accordance with the specifications outlined in Annex C of ISO UNI EN 3382-1:2009, it is imperative that the acoustic conditions be conducive to ensuring that musicians can perceive their counterparts with optimal clarity, whilst the chamber must be equipped to facilitate an adequate acoustic response. In order to obtain an objective evaluation of such a nuanced perceptual condition, it has proven useful to perform measurements on the orchestra platform with the sound source and microphone positioned in close proximity. The measurement process enables the calculation of two acoustic quantities: early support and late support. The former corresponds to the subjective listening aspect of ensemble conditions, while the latter corresponds to the perceived reverberation. [22,23]

It is reasonable to hypothesise that musicians adapt their performance techniques in response to the acoustics of the concert hall, given the significant impact of the acoustic environment on the auditory experience of music for both performers and listeners. [23]

Initial empirical evidence emerged from a laboratory experiment in which pianists performed in rooms with varying reverberation times. The results indicated a reduction in produced loudness in more reverberant environments, whereas the widest dynamic range was employed under moderately reverberant conditions. [24] Further more a recent investigation conducted in an anechoic room and simulating different acoustical enviroments indicated that the playing tempo was reduced not only in very reverberant rooms but also under anechoic conditions. [25]

Despite the fact that the space under scrutiny in this thesis is a recording studio as opposed to a conventional performance stage, the concept of stage acoustics remains pertinent. Indeed, it is vital for musicians to be able to deliver their best performance under optimal acoustic conditions, just as they would on the most significant stage of their careers.

Stage acoustics is concerned with the appreciation of the aocustics by musicians in halls and other performance spaces and various objective parameters have been introduced to describe it.

These parameters have been deduced by various empirical on different opera houses, theatres and other venues for performing arts.

The most widely recognized studies on objective measurements of stage acoustics are those proposed by Gade.[23]

Two different parameters from these measurements can be derived.

5.3.1.1 Early support

The ST_{early} was designed to quantify how early reflections support a musician's ability to hear their own instrument. When measured at a distance of 1 m, it represents mathematically, this parameters is the ratio of the reflected sound level within the 20-100 ms time interval after the arrival of the direct sound, at the numerator, and the level of the direct sound plus floor reflection measured in the time interval 0-10ms of the RIR, at the denominator.

5.3.2 Stage acoustic's measurements set up

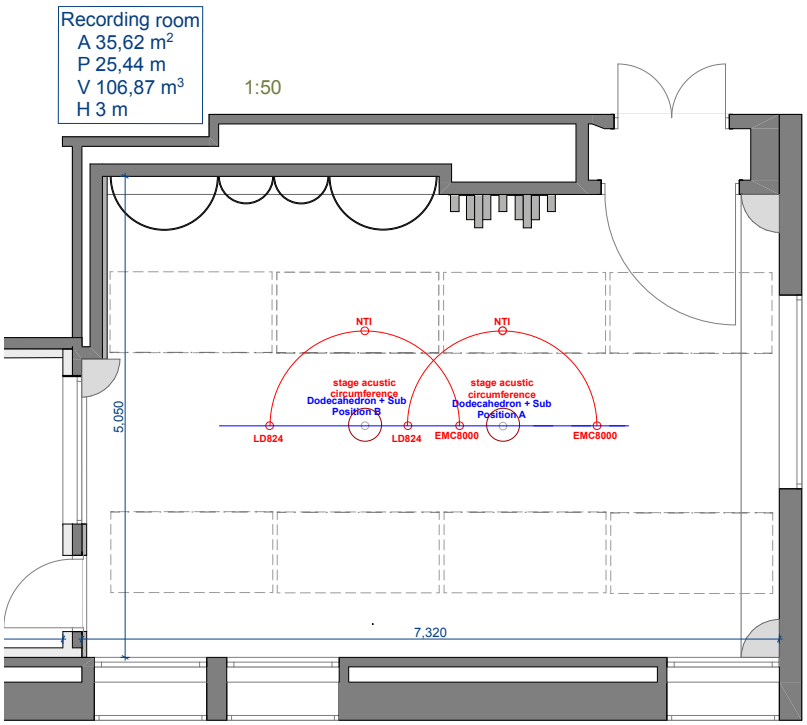


Figure 5.11 floorplan of the SERMIG's live room with the measruements point for stage acoustic

In accordance with the stipulations set out in Annex C of ISO UNI EN 3382-1:2009 [15], the measurement of the impulse response is to be conducted at a distance of one metre from the dodecahedral source. The microphone, for this purpose, is to be positioned at a height equivalent to that of the source, that is to say, either one metre or one and a half metres from the floor. It is also necessary to ensure that all reflective surfaces, with the exception of the floor, are maintained at a distance of more than 2 metres, while the edge of the stage must be no less than 4 metres away. It is imperative that measurements are taken from at least three distinct positions, with the source and the microphone being measured in each case. In the case of an orchestral configuration, chairs and music stands should be left in place, with only those located within a radius of 2 metres from the measurement point being removed.

In order to process the data, it is necessary to open the Aurora Acoustical Parameters plug-in and select the "Compute Stage Parameters" option. This will enable the calculation of indicators such as ST_1 and ST_{early}. Distinguishing between ST_2 and ST_{late} is also imperative, as these two phenomena may appear similar but are in fact distinct.

The measurement of stage acoustic parameters was achieved by positioning the microphones at a distance of 100 centimetres from the acoustic centre of the source, which was located at position A (Figure 5.11).

The SA Mic1 NTI stage position, with NTI microphone and source in position A, is located 290 cm from the control room wall along the longitudinal axis (i.e. 145 cm from point M1) and 100 cm from the source along the short axis. The height of the object in question, measured along the z-axis, is 1.5 metres from the floor, which is equivalent to the height of the source. The SA FON stage position, with LD824 sound level meter, is located 390 centimetres from the control room wall along the longitudinal axis (once more 145 centimetres from point M1) and 100 centimetres from the source. Along the short axis, the distance from the window wall is 243 centimetres. It is important to note that the height has been set at 1.5 metres, which is equivalent to that of the source.

The SA Mic2 BER stage position, equipped with a Behringer EMC8000 microphone, is located at a distance of 190 centimetres from the control room wall along the longitudinal axis (145 centimetres from point M1) and 100 centimetres from the source. On the short axis, the distance from the window wall is 243 centimetres, and the vertical height is 1.5 metres from the floor.

For the source located at position B (M3 centre), the microphone positions utilised for the measurement of stage acoustic parameters were situated 100 centimetres from the acoustic centre of the source.

The SB Mic1 NTI stage position, equipped with an NTI microphone and with the source in position B, is located 435 cm from the control room wall along the longitudinal axis and 100 cm from the source along the transverse axis. The measurement height is set at 1.5 metres from the floor, thus ensuring congruence with the height of the source.

The SB FON stage position, at which the LD824 sound level meter is utilised, is situated at a distance of 535 centimetres from the control room wall along the longitudinal axis and 243 centimetres from the window wall along the short axis. In this instance, the microphone is positioned at a height of 1.5 metres, which is equivalent to that of the source.

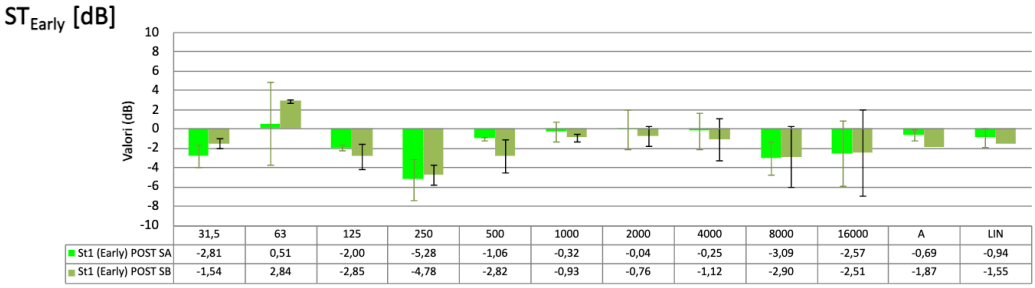
The BER stage position of the SB Mic2, utilising a Behringer EMC8000 microphone with the source positioned in location B, is situated at a distance of 335 centimetres from the control room wall along the longitudinal axis and 100 centimetres from the source. The distance from the window wall is 243 centimetres, while the height of the microphone remains fixed at 1.5 metres from the floor.

5.3.3 Results

As demonstrated in Figure 5.12, microphones A and B exhibit comparable levels of average linearity; nevertheless, B is observed to offer marginally superior support for ST_{early}.

Although the LIN is -0.94 dB for A and -1.55 dB for B, which confirms the aforementioned findings, the key frequency bands demonstrate a more nuanced reading: It is evident that B is considerably higher at 63 Hz, reaching up to +2.84, while A is less negative between 500 Hz and 4000 Hz.

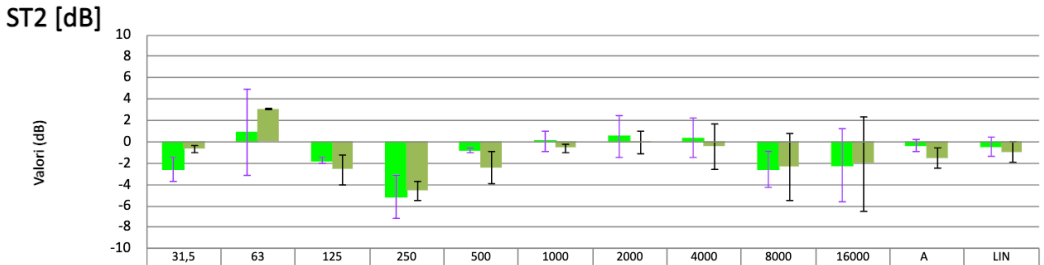
This finding indicates that, rather than a change in the total amount of early support, there has been a change in the spectral distribution. It can thus be concluded that setup B appears to generate greater body and support in the low frequencies. Conversely, setup A seems to offer slightly more balance in the midrange, and thus perhaps slightly more coherence in terms of the overall sound.



	31,5	63	125	250	500	1000	2000	4000	8000	16000	A	LIN
St1 (Early) POST SA	-2,81	0,51	-2,00	-5,28	-1,06	-0,32	-0,04	-0,25	-3,09	-2,57	-0,69	-0,94
St1 (Early) POST SB	-1,54	2,84	-2,85	-4,78	-2,82	-0,93	-0,76	-1,12	-2,90	-2,51	-1,87	-1,55

Figure 5.12 average ST_{early} of the three receivers for source position A and B

As demonstrated in Figure 5.13, even for ST_2, Setup B exhibits a lower limit of -1.00 dB compared to -0.45 dB for Setup A. This indicates that Setup A provides marginally superior room support.



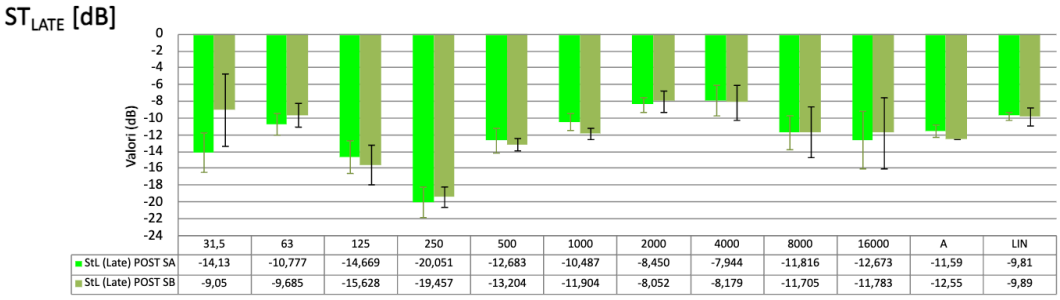
	31,5	63	125	250	500	1000	2000	4000	8000	16000	A	LIN
St2 POST SA	-2,59	0,89	-1,77	-5,11	-0,77	0,06	0,50	0,37	-2,58	-2,19	-0,35	-0,45
St2 POST SB	-0,71	3,07	-2,63	-4,62	-2,42	-0,62	-0,06	-0,41	-2,41	-2,06	-1,52	-1,00

Figure 5.13 average ST_2 of the three receivers for source position A and B

As demonstrated in Figure 5.14, the mean linear values of ST_{late} for the two configurations are almost indistinguishable, and even the A-weighted value – which is calculated to account for human hearing sensitivity, differs by less than half a decibel between the two configurations.

The fundamental conclusion is that the room's late-reflection component remains essentially stable, irrespective of the microphone setup or positioning that is considered.

This is of interest as it suggests that the hall exhibits a distinctive "signature" of late return that remains relatively constant despite minor alterations to the layout. This finding is consistent with the literature, which views ST_{late} as a parameter more closely linked to the broader field and less sensitive to local variations in early support.

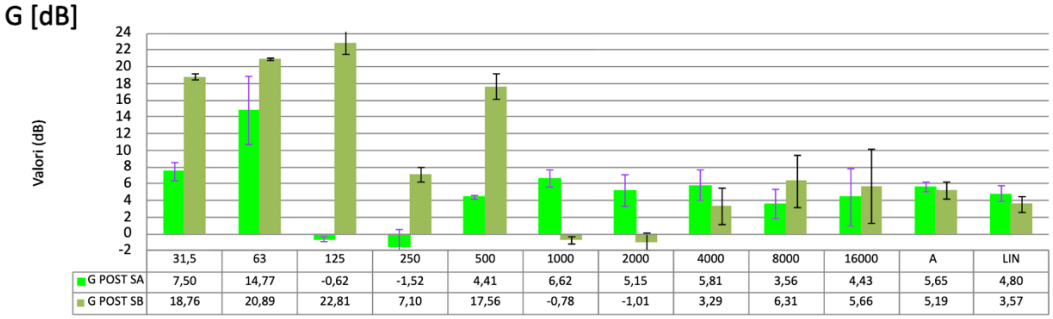


	31,5	63	125	250	500	1000	2000	4000	8000	16000	A	LIN
StL (Late) POST SA	-14,13	-10,777	-14,669	-20,051	-12,683	-10,487	-8,450	-7,944	-11,816	-12,673	-11,59	-9,81
StL (Late) POST SB	-9,05	-9,685	-15,628	-19,457	-13,204	-11,904	-8,052	-8,179	-11,705	-11,783	-12,55	-9,89

Figure 5.14 average ST_{late} of the three receivers for source position A and B

The observed combination of moderate ST_{early} values and strongly reduced ST_{late} values suggests that the room provides useful early acoustic feedback while limiting the build-up of late reflected energy. Such behaviour is consistent with environments where diffusive surfaces redistribute sound energy spatially without generating strong late reflections [20,26]. In musician-oriented recording spaces, this balance between early support and controlled late energy can facilitate ensemble interaction while maintaining acoustic clarity. [27,28]

5.3.4 Results of G



	31,5	63	125	250	500	1000	2000	4000	8000	16000	A	LIN
G POST SA	7,50	14,77	-0,62	-1,52	4,41	6,62	5,15	5,81	3,56	4,43	5,65	4,80
G POST SB	18,76	20,89	22,81	7,10	17,56	-0,78	-1,01	3,29	6,31	5,66	5,19	3,57

Figure 5.15 G average values obtained with the stage acoustic measurements setup A and B

These values of G, averaged over the three receivers used for the two setup of the source, calibrated in the anechoic chamber as preciously explained, stand for the acoustica gain of the room response, namely the energetic contribution of the ambient.

The first peculiarity that can be noted, observing Figure 5.15, is the behavioural differences between A and B in the low frequnecy band. producono un comportamento molto diverso alle basse frequenze.

In configuration B, the sound strength values G are elevated to a greater extent at low frequencies, with a pronounced increase observed below 500 Hz. In configuration A, the response is characterised by a more muted nature and a more uniform distribution across the spectrum. In contrast, in the mid-frequencies (1000–4000 Hz), configuration A maintains GG, whereas in configuration B, a more pronounced reduction in the parameter is observed. It is evident that the parameter G_G exclusively describes the total amount of sound energy reflected by the environment. Consequently, its interpretation must necessarily be supplemented with stage acoustics parameters, in particular ST_{early} and ST_{late}. It has been demonstrated that values of early support which are excessively high can result in the perception of the room being overly dry. Conversely, levels of late support which are excessively high are instead associated with acoustic conditions which are perceived as overly reverberant. [26]

In the case under examination, the joint analysis of the parameters indicates that the room exhibits a notable level of reflected energy, exhibited in a temporally regulated manner, with an early component that is sufficient to ensure adequate self-hearing for the musician and to facilitate ensemble interaction. This interpretation is also consistent with the results obtained from the room's standard acoustic parameters: the reverberation times indicate an acoustically "lively" environment, though not excessively reverberant, while the high values for clarity C_{80} and definition D_{50} suggest good intelligibility and effective signal control for both the musician and the sound engineer. Analysis using stage acoustics

parameters, however, allows for an additional level of interpretation, describing more directly the acoustic conditions perceived by the performer in the space. [26,29,30]

5.4 O-Zylia array with Insta 360 recording for VR integration

The measurements obtained from the O-Zylia microphone and the Instax 360 camera will be utilised for the auralisation of the sound environment. In order to complete the process, it is necessary that these recordings are accompanied by a monaural IR. This will be created with the LD824 capsule, in order to obtain point values that can be used for performing convolution within the software.
The measurement of acoustic parameters is facilitated by five microphone positions. The four aforementioned objects are located along the central longitudinal axis of the room, with a spacing of 145 centimetres between each. The remaining two, with a centre-to-centre distance of 100 centimetres, are situated on a transverse axis that intersects perpendicularly with the first at the second position, counting from the control room. With regard to the z-axis, the microphone is positioned at a height of 1.5 meters from the floor. This corresponds to the estimated average height of a person entering the room and assessing the soundscape of the recording space. In order to produce 3D audio-visual evidence of the room, an Instax 360 camera was utilised in the precise locations at which the O-zylia response was recorded.
In details, as you can see in Figure 5.9, *Position M1* is far away from the control room on X-axis (long axis) 145 centimetres, while on the Y-axis (short axis) is positioned at a distance of 243 centimetres from the window wall. The Z-axis (vertical axis) is positioned at a distance of 1.5 metres from the floor; *Position 2* is far away from the control room on X-axis (long axis) 290 centimetres, 145 centimeters from *Position 1,* whit same Y-axis (short axis) and Z-axis (vertical axis) position; *Position 3* is far away from the control room on X-axis (long axis) 435 centimetres, 145 centimeters from *Position 2,* whit same Y-axis (short axis) and Z-axis (vertical axis) position; *Position 4* is far away from the control room on X-axis (long axis) 580 centimetres, 145 centimeters from *Position 3,* whit same Y-axis (short axis) and Z-axis (vertical axis) position. Starting from *Positions 2*, the position *right* and *left* are each one 1 metre from the axis of the middle room, with the same Z-axis position of the previous measurement points.
The subwoofer and the dodechedron spherical source are positioned in *Position C*, which is angular and used for monaural impulses and for O-Zylia, is located 75 centimetres from the control room wall and 75 centimetres from the window wall, with a height of 1.5 metres from the floor.

5.4.1 O-Zylia positioning, recording and data exporting

In order to facilitate the process of sound acquisition using the ZYLIA ZM-1 microphone, it was first necessary to establish the software environment dedicated to interfacing and controlling the device. The installation of the primary components is the subsequent stage of the process. The following components are required for the functioning of the ZYLIA Studio: firstly, the microphone drivers and firmware; secondly, the ZYLIA Ambisonics Converter or ZYLIA Studio PRO, where this is deemed necessary. ZYLIA Studio functions as the primary management environment for the microphone, while the drivers

and firmware ensure effective communication between the hardware and the computer. The utilisation of optional tools is imperative in the conversion of raw 19-channel signals into first-, second-, or third-order ambisonic formats. These formats can then be employed in subsequent stages of analysis or three-dimensional rendering of the sound field. Following the installation of the requisite software, the ZM-1 microphone was connected to the computer via the dedicated USB port, which simultaneously provides power and data transfer. The correct connection is indicated by the front LED lighting up blue when the system is ready for use. Subsequently, it was possible to verify device recognition and access information about the model, connection status, and firmware version using the Zylia Control Panel.

Prior to commencing the recording procedure, meticulous attention was devoted to the orientation of the microphone, an aspect frequently underestimated yet of paramount importance for the accurate reconstruction of the sound field. The ZYLIA ZM-1 is characterised by a spherical geometry, comprising 19 capsules arranged in three concentric rings, with the Zylia logo, located on one side of the spherical surface, serving as an essential physical reference point. This logo is intended to identify the front axis of the microphone, and as such, it must be oriented with absolute precision towards the sound scene of interest or towards the direction that is to be defined as "front" in the subsequent ambisonic decoding. Incorrect orientation has the potential to compromise the spatial coherence of the recording, alter directional perception, and introduce errors in the reconstruction of the sound field during the rendering phase. In the context of the aforementioned recording, the microphone was placed on a stable support, with its height adjusted to 1.5 m in relation to the source. The logo was orientated towards the control room, with the vocal booth positioned behind it. The implementation of this procedure resulted in the establishment of a consistent reference axis between the recording and post-production phases, a prerequisite for subsequent ambisonic analysis. The position of the subject was established within the software program and remained constant until the conclusion of the recording sessions.

Subsequent to the definition of the orientation, the selection was made in the Control Panel for RAW 19 Channels mode, a prerequisite for the capture of the independent signal from each capsule. The settings were configured with a sampling frequency of 48 kHz and a depth of 24 bits, parameters that guarantee an optimal balance between sound quality and data management. The subsequent stage of the process entailed the verification of the input levels. Utilising the real-time metering display for each capsule, the gain was calibrated to ensure that peak levels remained within the range of -12 dB to -6 dB. This procedure was undertaken to guarantee sufficient headroom and to avert the occurrence of saturation.

The resulting multichannel WAV file contained the 19 raw signals, which were then processed using MATLAB.

5.4.2 O-Zylia data analysis

In the MATLAB environment, the analysis procedure is performed by extracting the directional components from the multichannel signals acquired with the microphone array. Each impulse response recorded with O-Zylia is transformed into a 16-channel file, corresponding to the third-order ambisonic format (Figure 5.16). In practice, the system captures the three-dimensional sound field around the microphone and encodes it into ambisonic coefficients up to the third degree, allowing for a much more detailed spatial representation than lower-order formats. This facilitates highly accurate reconstruction of

the directionality and depth of the acoustic environment, rendering RIRs suitable for immersive applications, virtual reality, augmented reality, and advanced binaural rendering.

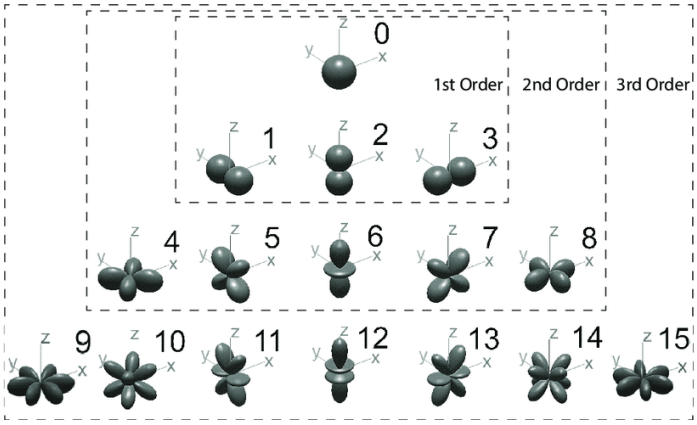


Figure 5.16 Ambisonics formats up to third order graphically explained

Subsequent to this step, the script verifies the sampling rate of the recordings, which in this case is 48 kHz, with a view to ensuring that all data is consistent and correctly aligned. The subsequent step in the procedure requires the user to specify the time intervals, measured in seconds, within which the analysis of the sound decay behaviour is to be conducted. This step is of crucial importance, exspecially for the Time Decay Function diagram.

5.4.3 Time Decay Function diagram

The Time Decay Function (TDF) diagram facilitates the isolation of precise portions of the impulse response, thereby enabling the focus to be directed on the segments in which the reverberation develops and decays. Two analysis modes can be selected, from the MATLAB script, which are applied to the selected time intervals. Forward analysis examines the behaviour of sound decay within the specified interval, thus focusing on the portion of RIR between the chosen time limits. Conversely, backward analysis involves the examination of the decay behaviour that occurs outside the specified interval. This analysis considers all sound information that has been excluded from the defined time window. This enables a direct comparison of the two perspectives, thereby facilitating a more comprehensive understanding of the evolution of reverberant decay. Conventional acoustic parameters primarily focus on decay; however, this method also examines the development of the sound spectrum in the initial stages of RIR, namely the direct sound and the first reflections, as these are pivotal to comprehending how we perceive the spatial and auditory qualities of our environment, in terms of clarity of the signal and the perception of source position too.[28]

In accordance with this method, prior to the calculation of the discrete Fourier transform (DFT), a rectangular window is employed. While it may not be the most pristine from a spectral perspective, it ensures consistency between windows of varying durations for the processing of impulse responses. One could elect to process the signal with Hann or Hamming windows, which have been shown to be smoother

and to produce fewer side lobes and leakage problems than a rectangular window. However, these windows cannot be used easily because they require fade-in and fade-out times, which would be inconsistent when analysing many windows of different durations. The responses to the impulse, once windowed in their respective time intervals, are transformed using the DFT. The absolute values of the results are then averaged across all sources. The rising trend of the curves demonstrates the contribution made by each additional time segment. The method enables the clear and effective visualisation of how the frequency response develops at the listening point as increasingly larger portions of the RIR are considered.[28]

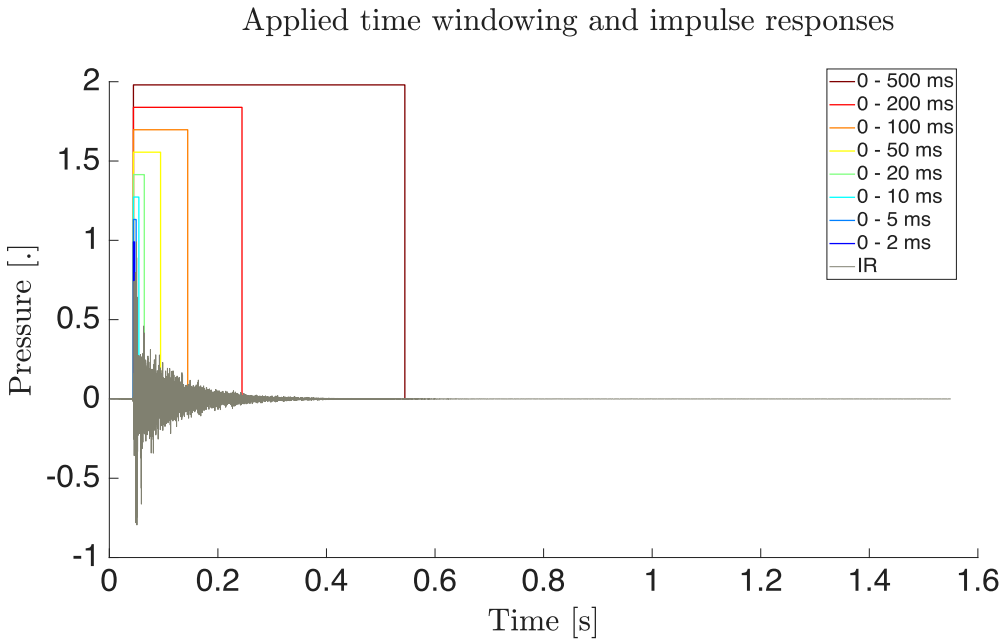


Figure 5.17 Pressure over frequency plotting of the decay of the IR

The proposed analysis, showed in Figure 5.17, has thus far exclusively considered the monaural frequency response, without providing any information regarding the direction of sound origin.

However, by employing conventional spatial audio techniques, the method can be expanded to encompass the directionality of the sound field in real-time to have a spatiotemporal analysis.

The graph shows the impulse response (IR), which is measured in the time domain. It also shows a series of time windows of different durations that have been applied to the signal. Using time-windowing allows specific portions of the IR to be isolated, enabling the contribution of the direct field, early reflections and reverberant component to be distinguished. As discussed in the literature, very short windows (2–10 ms) almost exclusively include the direct arrival, resulting in a more regular frequency response that is free from the comb-filtering effects produced by reflections.[3]While these windows are useful for source analysis or equalisation in pseudo-anechoic conditions, they result in reduced spectral resolution at low frequencies due to their limited duration. [31] Intermediate windows, with a typical range between 20 and 50 milliseconds (ms), encompass early reflections, thereby engendering a more undulating frequency response that is concomitantly more representative of the interaction between the source and the environment. It has been demonstrated that the inclusion of more extended windows, ranging from 100 to

500 milliseconds, results in the majority of the reverberant tail being captured. In such instances, the frequency response exhibits significant irregularities due to the presence of multiple interferences. However, the signal retains the essential information required for conducting an energy analysis of the decay process. This includes the estimation of the reverberation time in accordance with the standards outlined in ISO 3382-1:2009.[15] The selection of the time window has been demonstrated to exert a substantial influence on the stability of the frequency response and the interpretation of the acoustic phenomena associated with the measurement. Consequently, time-windowing is regarded as a pivotal instrument in the characterisation of impulse responses and the processing of measurement data. [32]

5.4.4 Rose decay diagram

There are several approaches to obtaining directional information from a spatial IR; in this case, the analysis is based on an unconstrained least squares solution, obtained from Time Difference of Arrival (TDOA) estimates. These estimates allow the origin of the sound to be determined by estimating the direction of incidence in the arrival times at the different capsules of the microphone array. This methodology enables the estimation of the direction of each individual audio sample of the impulse response, measured at the geometric centre of the 3D probe.
Consequently, in lieu of a rudimentary monaural RIR, the directional analysis engenders a spatial impulse response, replete with information pertaining to the genesis of the sound waves over time. This facilitates a considerably more comprehensive and authentic representation of the acoustic behaviour of the environment. [28]
In order to visualise this data, a rose plot is employed, a graphical representation in polar coordinates that intuitively demonstrates how a certain quantity, in this case sound energy or sound decay, is distributed in different directions. The diagram, reminiscent of a "wind rose," provides detailed information on how direct sound and reflections reach the microphone, depending on the arrangement and diffusion of surfaces in the environment. In practice, multichannel recordings obtained with a spherical array such as Zylia's, consisting of 19 capsules, allow the directional component of sound energy to be reconstructed. The decay of sound energy is then evaluated for each channel or group of channels, thereby generating a series of decay curves associated with specific directions in space. When these values are plotted on a polar diagram, the rose plot is produced, which makes it immediately apparent from which directions the sound persists longer or decays more quickly. This provides an intuitive and detailed understanding of the spatial distribution of the sound field in the environment.

For each directional component, the energy decay curve is calculated in the time domain, applying, where appropriate, the inverse energy integration method proposed by Schroeder (1965), a method which is widely used for determining reverberation times and derived acoustic parameters. [33]

$$h_l'(t|\hat{\theta}_l(t), \hat{\varphi}_l(t)) \triangleq \left[h(t)_l, \hat{\theta}_l(t), \hat{\varphi}_l(t)\right] = DIR\,\{h_l(t)\} \qquad (6.2)$$

Instead of having a monoaural impulse response, the directional analysis provides spatial information for the room impulse response as in equation (6.2) where $\hat{\theta}$ and $\hat{\varphi}$ denote the azimuth and elevation angle estimates respectively. [28] In order to facilitate the interpretation of the results, the sound field is divided into three primary sections: lateral, transverse, and median.
The median section is indicative of the frontal direction, i.e. the principal axis of listening or recording. For instance, it corresponds to the direct sound emanating from the source positioned in front of the array.

The lateral section comprises the directions to the sides of the array, i.e. the sounds emanating from the right or left. This component is of significant importance in the field of acoustics, as it contributes to the perception of spaciousness and sound envelopment.

Finally, the transverse section is defined as referring to intermediate directions that are neither entirely frontal nor fully lateral. It can be posited that these directions represent the 'crosswise' direction with respect to the main axis.
Graphically speaking it can be clearer if we explain this tree different sections treating our O-zylia as a Head and Torso Simulator – HATS: the blue represents the transverse section (Figure 5.20), the green one stands for the median section (Figure 5.19), while the red one for the lateral one (Figure 5.18).
This methodology aligns with the procedures outlined in ISO 3382-1:2009, [8] which pertains to the measurement of room acoustic parameters. It facilitates the integration of temporal analysis with spatial analysis, thereby offering a more comprehensive description of the sound field.

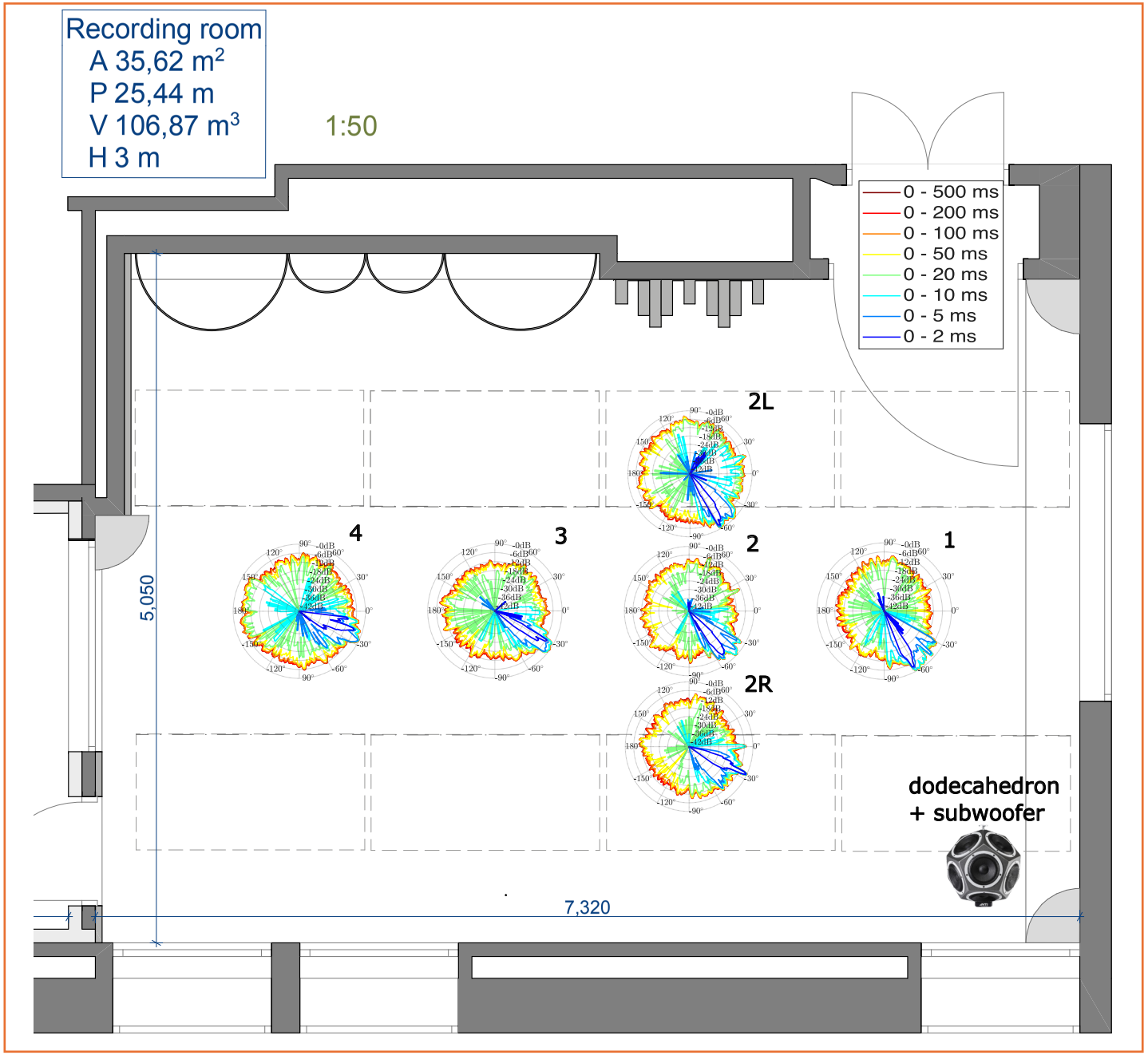


Figure 5.18lateral section of the decay roses

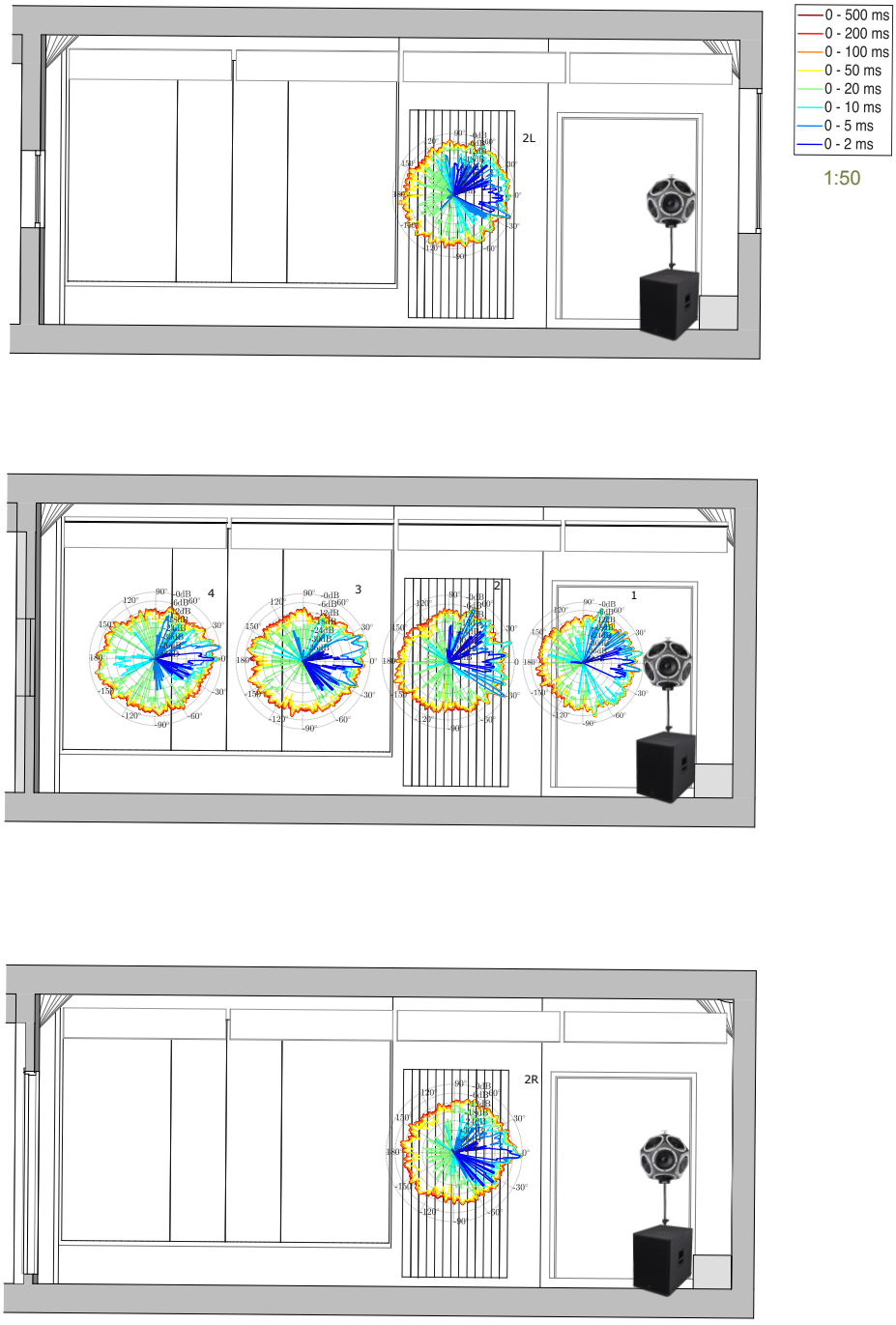


Figure 5.19 median section of the rose decay

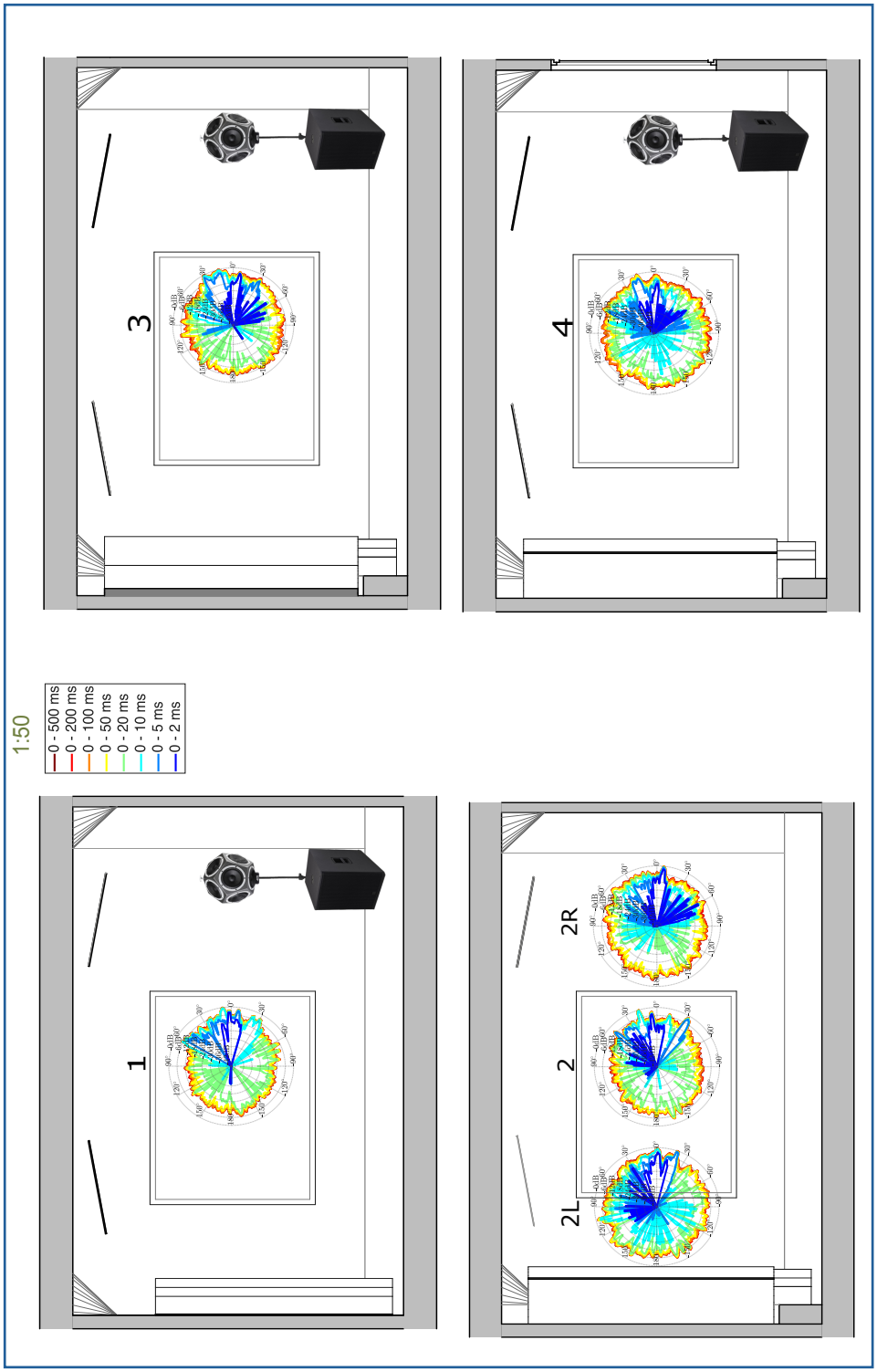


Figure 5.20 transverse section of the decay roses

5.4.4.1 Frequency behavioural of sound analisys

In addition to time domain graphs and rose plots, the analysis of the recordings performed with MATLAB provides us with graphs to analyse the frequency behaviour of sound within this environment. The following graphs illustrate the evolution of the frequency response as a function of the time window employed for the integration of the impulse signal. Each curve illustrates the variation in level in decibels (dB), considering solely the decay of the impulse within a specified time interval, ranging from 0–2 milliseconds (ms) to 0–500 milliseconds (ms), as the frequency varies. This method of representation is widely employed in the field of acoustic literature [3,34], as it facilitates the evaluation of the direct component, early reflections, and reverberated field as discrete entities within the overall frequency response. This representation elucidates the function of the various temporal segments of the impulse, thereby elucidating the manner in which each contribution effects a modification of the measured spectrum. The curves obtained with very short windows (e.g., 0–2 ms, 0–5 ms, or 0–10 ms), which are generally represented in blue and light blue, essentially include only the component. In the absence of significant reflections, the response appears to be smoother, with minimal oscillations and lower overall levels. This tendency is consistent with the behaviour of the loudspeaker in near-anechoic conditions, as demonstrated in numerous studies on the direct response of loudspeaker systems[35].

As the duration of the time window increases (0–20 ms, 0–50 ms, 0–100 ms), represented in the green, yellow, and orange curves, early reflections and the first resonances of the environment are taken into account. It is evident from the behaviour of this segment of the signal that the overall level and the emergence of more pronounced undulations can be observed. These undulations are a consequence of constructive and destructive interference between the direct signal and the reflected contributions. This phenomenon has also been described by Farina in the context of impulse response analysis with moving window techniques [36]. This effect is particularly evident in the mid-high frequency bands, where nearby reflections introduce more detailed variations in the spectral shape.

The curve relating to the longest window (0–500 ms), shown in dark red, represents the complete frequency response of the environment, including both early reflections and the reverberated component. The figure is indicative of the maximum density of peaks and troughs, thereby highlighting the combined effect of multiple reflections and the energy that remains in the sound field.

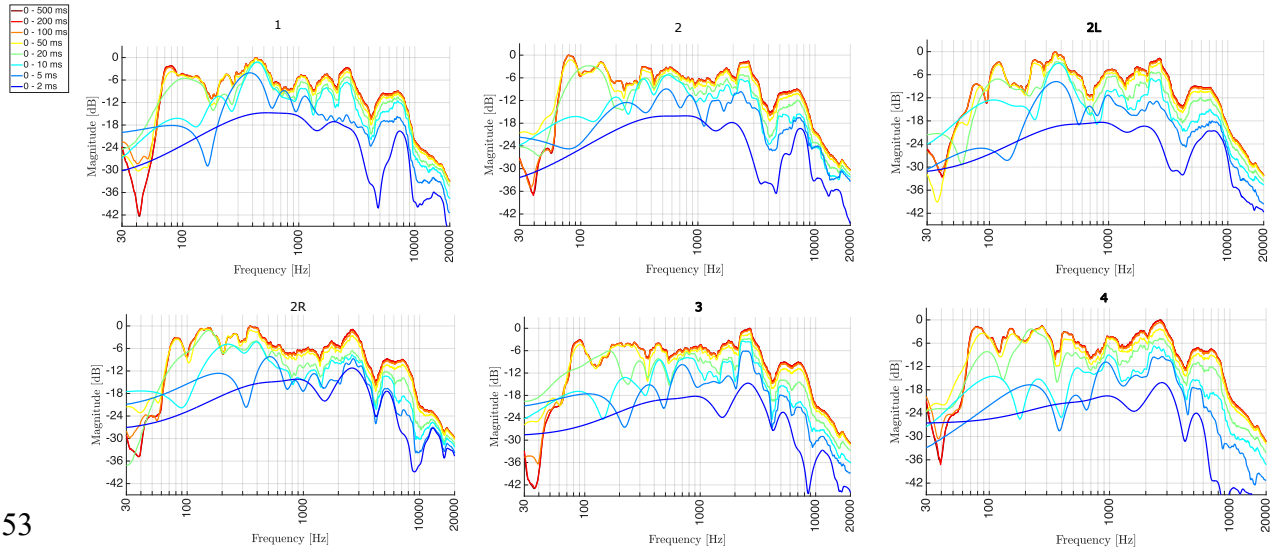


Figure 5.21 Frequency response of time-windowed energy decay (0-2 ms to 0-500 ms) at receiver positions 1-4 (including lateral points 2L and 2R), showing the spectral contribution of early and late reflections across the 30 Hz-20 kHz range.

This representation is advantageous in that it enables the distinction between the system's contribution and the response of the environment, thereby facilitating the analysis of reflections, loudspeaker-room interaction, and the effect of reverberation on the final frequency response. Its usefulness is recognised in numerous contexts, ranging from the acoustic design of environments to the optimisation of sound diffusion systems [3,35]. A preliminary comprehensive evaluation of all measurement points has been conducted, revealing a substantial discrepancy between 2 ms and 500 ms windows. This finding suggests that the various frequency bands are significantly influenced by the environment's response and the presence of stationary modes.

It is hypothesised that certain strong imbalances may be attributable to phase cancellations of the various directional reflections of the sound source, in conjunction with cancellations that may be induced by the 19-channel microphone array.

With regard to the low frequency range, the response level is found to be minimal, and the curve is observed to be relatively smooth, exhibiting a slight overall attenuation from 500 Hz onwards. This attenuation is more pronounced in the central area of the room, in the vicinity of the control room. Indeed, at the third and fourth measurement points, the frequency response of the direct sound components is more linear in this frequency range.

Within the broader time frames, the reflections of this frequency range exhibit a substantial increase, manifesting as dips that approach the maximum scale (0 dB) in specific measurement positions, particularly between 100 Hz and 200 Hz. This behaviour is indicative of energy accumulation and is reflective of the prevailing environmental conditions. With regard to the auditory aspect, the acoustic characteristics of the listening chamber are characterised by a relatively warm sound, albeit with a certain degree of heterogeneity. However, the absence of frequencies below 80 Hz results in a certain lack of depth.

Conversely, for mid-frequencies, the response curves exhibit a propensity to overlap more rapidly, even within time windows spanning from 20 to 50 milliseconds. It is evident that there is a particularly notable uniform balance throughout the room, with moderate dips, especially between 400 Hz and 1 kHz. It is evident that reflections with times less than 20 milliseconds can exert a significant influence on the system, giving rise to comb filtering phenomena and oscillations in the response. These fluctuations in the responses can be attributed to copies of the direct sound that are delayed by a few milliseconds. At the perceptual level, this could result in subjective colouration, depending on the delay time and the height of the relative masses. [3]

In contrast, when very short windows (2–5 milliseconds) are taken into consideration, the direct response appears to be significantly smoother, thereby highlighting the linear behaviour of the system.

In the high frequency range up to approximately 3–4 kHz, the wavelengths are sufficiently greater than the array dimensions and the inter-capsule distance, allowing for a coherent representation of the sound field. In this frequency range, the frequency responses obtained using short and long time windows can also be considered reliable from a timbral point of view, as they realistically reflect the contribution of direct sound and early reflections, in accordance with the principles of time-frequency analysis applied to acoustics [37].

It is imperative to exercise greater caution when interpreting the graphs above 4 kHz. In this spectral region, the wavelength becomes comparable to the size of the array, resulting in spatial aliasing phenomena, interference between capsules, and directional filtering. These effects are comprehensively documented in the field of spatial sampling theory [38]. The observed effects manifest as spectral ripples, marked differences between channels, and a progressive roll-off of the high-frequency response. These effects cannot be attributed to the acoustics of the environment. It is imperative to note that any irregularities observed above 8–10 kHz should be interpreted primarily as a consequence of the intrinsic limitations of the measurement system, rather than as actual defects in the acoustic response of the system under analysis.

It is noteworthy that certain notches observed in these graphs may be attributable to 'dip seat' effects, a phenomenon attributable to the characteristics of the surfaces in the room. For instance, the measurements obtained at positions three and four were in close proximity to a grand piano, a substantial reflective surface with the potential to substantially influence the outcomes.

6 Audio simulation for Virtual Reality

6.1 Sound simulation models

In order to elaborate on the methodology employed in this project, which entailed simulating the acoustic behaviour of the recording studio located within the Arsenale della Pace in Turin, it is imperative to introduce the overarching concept of sound propagation simulation, which can be defined as a vibrational field. Dr. Michael Vorländer is a pioneering figure in the field of realistic sound simulation in virtual environments. He is a highly regarded researcher and professor of technical acoustics at RWTH Aachen University in Germany, where he served as the director of the Institute of Technical Acoustics (ITA), a leading research centre in this domain. His contributions have elevated acoustic simulation to the status of a fully-fledged engineering discipline, thereby demonstrating the feasibility of synthesising the manner in which an environment responds to sound based on physical and mathematical models. His book *Auralization: Fundamentals of Acoustics, Modeling, Simulation, Algorithms, and Acoustic Virtual Reality* is widely regarded as a foundational text in the field.

This text undertakes an analysis of the scientific and technical foundations of auralization, defined as the process of generating audible sound simulations from physical data. The author commences their investigation by exploring the physical behaviour of sound and vibration waves, describing their interaction with enclosed spaces through phenomena such as reflection, absorption, and diffusion. A substantial portion of the text is dedicated to psychoacoustics, wherein the author employs a contemporary perspective on the study of acoustics to explore the human ear's perception of sound intensity and direction in spatial environments. The text further illustrates mathematical models and signal processing algorithms for digitally simulating acoustic behaviour. The work under consideration here addresses the use of CAD modelling methodologies and ray tracing as a means of predicting sound quality in architectural environments such as concert halls. [39]

6.1.1 Mathematical models for sound simulation

The objective of the simulation is to determine the propagation of sound and vibrations. These phenomena are often described using Green's functions or transfer functions, which represent the response of the filter system between a source and a receiver.
Green's functions are derived from the formulation of the wave equation, which is derived from the potential function, $g(r|r_0)$. The latter is contingent on r_o, the position of the source, and r, the position of the receiver, and is a potential function from which the physical quantities of the sound field are obtained through derivation operations. It has been established that the sound pressure p is in fact linked to the time derivative of g, while the particle velocity v is linked to the spatial gradient of g.

This description of the system's response is especially pertinent in the context of a point source with its own volume flow, Q, which undergoes a loss of pressure as the distance increases, accompanied by a time delay due to the spatial separation between the source and the receiver.

It is evident that, given our presence within a closed environment rather than a vacuum, Vorländer's research demonstrates that Green's formulation gives rise to an integral, designated as Helmholtz-Huygens. This integral facilitates the summation of the effects emanating from sources distributed within the volume, in addition to the effects of reflections occurring on surfaces.
Utilising this line of reasoning, he posits the assertion that "the integration surface can also serve as a virtual surface, wherein the sound field undergoes expansion into elementary (secondary) sources" (Huygens' principle). It is noteworthy that the surface source configuration consists of monopoles and dipoles. This methodology is employed to illustrate the phenomenon of sound reflections.

The Green's function is defined as the response of the system when the time domain is considered and an impulsive point source is employed, thereby yielding the solution to the wave equation when the source is a Dirac delta. It is imperative to note that Green's functions are filters that modify the behaviour of the signal on its path from the source to the receiving point. Indeed, the room constitutes an LTI (Linear Time-Invariant) system in signal processing, whereby g is the audio signal with which it must be convolved to facilitate comprehension of its behaviour within the environment. It is from this that the system's transfer function will be obtained.

The concept of reciprocity further assists in clarifying this phenomenon by stating that, within an LTI system, the transfer function, and consequently the Green's function, remains unchanged if the order of the audio signal's path is reversed. Indeed, inverting the positions of the source and receiver does not modify the manner in which the environment filters the audio signal, i.e., it does not affect the transfer function of the room.

In conclusion, it is imperative to acknowledge that irrespective of the method employed, the objective is to identify a filter that facilitates the transmission of the signal from the source to the receiver.[39]

6.1.2 Frequency domain model

Wave-base models are a class of simulation methods that solve the Helmholtz equation. These models are particularly well-suited to low frequencies, where the wave nature of sound is dominant. [39]

- **FEM – Finite Element Method**: this is the basis of this theory is the discretisation of the entire volume of the acoustic field into a multitude of small elements. Utilising a variational approach, known as Hamilton's principle or the principle of least energy, the objective is to identify the field quantities, which in the case of sound are sound pressures, that minimise the total energy of the system. The total energy of the system is calculated as the sum of the energies of all the individual elements into which the volume is divided.
  The acoustic problem is translated into a matrix system, which is then solved using either a direct method to obtain the sound pressure, or an indirect or modal method, which returns the eigenvalues corresponding to the resonance frequencies and modes of the room.
  This approach is regarded as the most rigorous method for studying modal behaviour and low-

frequency resonances in closed volumes, based on the variational principle of energy and volumetric discretization.[39]

- **BEM – Boundary Element Method**: it represents the predominant alternative to the FEM, particularly in scenarios involving outward radiation or scattering. Whilst the finite element method (FEM) divides the total air volume into a multitude of three-dimensional elements, the boundary element method (BEM) only discretises surfaces. The BEM is predicated on the Helmholtz-Kirchhoff integral, which employs Green's functions for free space, thereby delineating the propagation of sound from a point source in an unobstructed space. This method has been demonstrated to be a valuable tool for addressing open-ended problems, primarily due to its capacity to manage infinite domains while discretising solely the surface of the object. However, it should be noted that at elevated frequencies, this approach can incur a substantial computational burden. [39]
- **SEA – Statistical Energy Analysis** : SEA represents a paradigm shift compared to FEM and BEM: rather than calculating the precise sound pressure at each point in space (deterministic approach), it calculates the average energy stored in entire subsystems (statistical approach). The SEA method was developed to address high-frequency problems, where methods such as FEM become impractical. Indeed, while the finite element method (FEM) struggles to create meshes with millions of tiny elements to represent high frequencies (small wavelengths) that would make the calculation complex and costly, the finite element analysis (FEA) is based on the fact that, at high modal density, the modes overlap statistically. This method is particularly useful for the vibro-acoustic analysis of complex high-frequency systems, such as buildings, ships, or vehicles, as it is based on balancing energy flows between coupled resonant subsystems, thereby providing reliable average values where deterministic methods would fail due to excessive sensitivity to geometric details. [39]

6.1.3 Time domain model

These methods simulate the propagation of the sound pulse over time and neglect the wave nature in order to focus on the energy propagation paths.

- Geometrical Acoustics (GA): is the cornerstone of acoustic simulation for rooms in the mid and high frequencies. This approach differs significantly from the wave-based (FEM/BEM) and statistical (SEA) methods previously discussed. It is based on the assumption that sound propagates in rays, analogous to geometric optics.
    - **Ray tracing model**: it is a stochastic method, whereby the source emits a large number of sound particles or rays in directions usually distributed over a sphere. It has been established that each particle travels at the speed of sound until such a time as it makes contact with a surface and is reflected, thereby losing a proportion of its energy. The receiver is perceived as a volume traversed by a ray. The duration of the traversal of the receiving volume is documented in a histogram, which is the response to the energy pulse of the room. This method is considered to be highly effective in the management of

diffuse reflection and the calculation of the reverberation tail. In scenarios where the number of rays is inadequate, the outcomes exhibit statistical variations.

- **Image source model**: it is a deterministic method based on the principle that a specular reflection on a wall can be modelled by a virtual source, also termed an image, placed behind the wall, symmetrical with respect to the original source. In order to verify the validity of the beam path, a rigorous geometric test is required. The temporal resolution offered is of a very high order, ideal for early reflections. However, the number of image sources increases exponentially with the order of reflection, making it impossible to calculate late reverberation. [39]

- **Radiosity** is a geometric method that models sound propagation as an exchange of energy between discrete elements of the room's surface. This established technique is widely utilised in the field of computer graphics for simulating lighting, and was subsequently adapted for application in acoustics. The fundamental principle underpinning this phenomenon is that each surface element is exposed to energy by radiation, both directly from the source and from its reflections due to other elements present in the room. Consequently, this element functions as a secondary source, radiating a proportion of the energy received back to the other elements. The physical assumption that underlies this mechanism is that all reflections on the walls are ideally diffused according to Lambert's Law of Cosines. This is in contrast to a mirror, which reflects in a single direction. Instead, the intensity of the diffused reflections is dependent on the cosine of the angle of emission and independent of the direction of arrival of the incident sound. This approach is regarded as the most rigorous method for modelling diffuse reflection and energy decay in environments where the walls are not mirroring but diffusing. [39]
- **Waveguides** is a theoretical representation of sound propagation. It posits that sound waves travel through connected tubes, with delays representing propagation time, connections representing reflections and transmissions, and filters representing losses that depend on frequency and energy.
  This method is regarded as the optimal approach for sound synthesis (musical instruments, voice) due to its efficacy in simulating 1D resonances. However, it can be expanded through the utilisation of mesh networks (Digital Waveguide Mesh) to simulate propagation in 2D and 3D environments, thereby serving as a conduit between physical modelling and digital signal processing. [39]

6.2 Simulating software of sound in rooms

The algorithms employed by the majority of programs are based on geometrical acoustic models. These models approximate the propagation of ray waves and their energy. It is evident that these algorithms neglect the wave nature of sound, incorporating diffraction and interference. This simplification is only valid if the dimensions of the room are large compared to the wavelength, i.e., above the Schroeder frequency. In the field of numerical analysis, a common approach to the simulation of electromagnetic scattering is to utilise a hybridisation of two classical geometric methodologies. The first of these is

stochastic ray tracing, otherwise referred to as the Monte Carlo method. This approach has been shown to be effective in the management of scattering and late reverberation, though it is limited in terms of temporal resolution. The second method is the image source model, which, while accurate for the initial part of the impulse, exhibits an exponential increase in computational cost with the order of reflection. Consequently, a hybrid use of both methods is frequently selected. Contemporary software such as Odeon, Catt, and Ease eschew a single, unadulterated method in favour of a combination of approaches. These utilise image sources for the initial reflections in order to achieve precise localisation and timbre, and ray tracing for reliable calculation of late and diffuse reverberation. [39]

For the audio simulation in this thesis it has been used Odeon, which is a specialized room acoustics simulation software based on geometrical acoustics, whereas AuVi is an open-source Blender-integrated framework that enables customizable acoustic simulations within a flexible 3D modeling environment.

6.2.1 Physical Constraints of Ray-Based Acoustic Modelling

As discussed in the scientific literature, methods based on geometrical acoustics, including ray tracing employed in the software used in this thesis, Odeon and AuVi Suite, introduce a series of intrinsic limitations related to the physical assumptions on which they are based. In this approach, sound propagation is modelled as a set of rays travelling in straight lines, implicitly assuming that the wavelength is negligible compared to the geometric dimensions of the environment.[40]
This simplification allows for considerable computational efficiency, but involves the exclusion of the main wave phenomena that characterize the actual propagation of sound.
These methods are based on a high-frequency approximation, and are therefore reliable above the Schroeder frequency, where modal overlap is sufficiently high to allow for a statistical description of the sound field.[40] In the case of values falling below this threshold, the low modal density becomes incompatible with the assumptions underpinning geometrical acoustics, resulting in a consequent reduction in the simulation's accuracy.[40] In particular, effects such as diffraction, interference and modal coupling are not represented in a physically rigorous manner. These effects derive from the solution of the wave equation and cannot emerge from a purely energetic or geometric model. [41]

Ray-based methods are typically employed to analyse sound energy quantities rather than acoustic pressure or particle velocity. They utilise simplified surface parameters, such as absorption and scattering coefficients, which do not incorporate phase information.[40] The result is a representation that describes the transport of sound energy, but not the acoustic field in its entirety. This suggests that the behaviour of sound in the presence of geometric discontinuities, impedance variations, or complex boundary conditions is treated in an approximate manner. In situations where propagation is influenced by wave phenomena – for example, in the vicinity of surfaces, edges, or non-ideal geometric configurations – ray tracing models fail to capture the mechanisms of interaction between the sound field and the environment. As highlighted in studies based on numerical methods, techniques such as ray tracing and image source explicitly neglect these effects, making wave-based methodologies necessary when a physically complete description of the acoustic field is desired. [41]

6.3 Calibration and validation of a model

In addition to comprehending the mathematical models employed in acoustic simulation for virtual reality environments, it is imperative to elucidate the central issue prior to delving into the parameterization of the software utilised in this thesis: the reconstruction of sound in a space within a virtual simulation that is both authentic and convincing.

A figure who has addressed this issue in a complementary manner to that of Vorländer's work is Brian F. G. Katz, a researcher at the Centre National de la Recherche Scientifique (CNRS) and the Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI). While Vorländer is a reference point for the reliability and engineering soundness of acoustic simulation, based on numerical models and predictive approaches, Katz shifts the focus to a perceptual and experiential dimension. Katz's position is that, in addition to the computational intricacies inherent in simulation software, a model can only be deemed authentic if it is deemed perceptually credible by the listener. In this perspective, the author highlights how virtual reality reconstructions have historically prioritised the visual component, often neglecting the sound component. [42]It is precisely from this observation that his contribution to the development of the concept of acoustic archaeology, understood as the study of the past through the reconstruction of the sound of spaces, emerges.

The method proposed by Katz is an operational and reproducible procedure for calibrating and validating a "virtual" acoustic model The fundamental concept is to regard the simulation as a "virtual measurement": a Room Impulse Response (RIR) is derived from the model, the acoustic parameters employed in actual measurements are extracted, and the model is calibrated until the discrepancies are perceptually negligible, with the Just Noticeable Difference (JND) threshold serving as the termination criterion. [42]

6.3.1 Reference data and comparison metrics

In the calibration process proposed by Katz, the starting point is represented by the actual acoustic impulse response (RIR) measurements taken within the physical environment under study. Katz's methodology involves the initial measurement of RIRs, which are then used to identify a set of perceptually relevant acoustic parameters. These parameters are then employed as calibration targets. These parameters are organised into two broad categories, each of which is associated with specific dimensions of the listening experience. The initial consideration pertains to the behaviour of reverberation and energy decay within the sound field, characterised by 20 dB reverberation time (T_{20}) and Early Decay Time (EDT), which serve as metrics for the rate of sound energy dissipation in the ambient environment. These parameters are closely related to the perception of spatiality and acoustic envelopment.

The second category is associated with sound clarity, and for this purpose, the clarity parameters C_{50}, which is more relevant for speech, and C_{80}, which is generally used in musical contexts, are utilised.

The comparison between simulation and reality is not made on a single global value, but by analysing these parameters as a function of frequency, typically on octave bands between 125 Hz and 4000 Hz. Moreover, the evaluation is conducted on a series of source-receiver configurations distributed in space. This approach is employed to both capture the average behaviour of the environment and its spatial variations. [42]

Katz adopts the baseline values from ISO 3382-1: 5% for reverberation parameters and 1 dB for clarity parameters, and proposes using them as thresholds for perceptual relevance. [15,42]

6.3.2 Methodological principle and procedure in 6 steps

A crucial practical consideration is that, despite the availability of certain software that provides the parameters directly, Katz employs a more meticulous approach by recalculating the parameters from the simulated Raster Inverse Distance (IR) files using the same analytical tool utilised for the measurements. This methodology is employed to circumvent any potential discrepancies that may arise from the implementation of the calculation algorithm (noise estimation, windowing, etc.). This approach is adopted to ensure that the simulation and actual measurement are analysed in a consistent manner. Furthermore, simulated RIRs are ideally "noise-free"; in order not to distort automatic noise detection algorithms, which are often designed for real measurements, Katz adds very low-level Gaussian white noise, approximately -65 dB, before analysing the parameters.
Katz proposes a general procedure applicable to GA that generate RIRs through auralization, and then outlines it in practical terms. The procedure can be summarized in six steps.

- **Step 1: Preparation and Preliminary Assignment**
    - Physical measurements: The measurement of the room impulse response (RIR) is undertaken in situ, with the results serving as a reference for calibration purposes.
    - Model creation: The creation of an unalterable geometric model is initiated, and preliminary acoustic properties are assigned to it.
    - Material research: Relevant absorption coefficients are identified from databases to define a reasonable range of values to be used during the process.
- **Step 2: Analysis of Variance (Run-to-run variation)**
    - Repeated simulations: Several repetitions of the initial simulation are performed (e.g., 10 times).
    - Quantification of uncertainty: the standard deviation is averaged for each acoustic parameter in order to quantify the intrinsic variation of the software due to the stochastic implementation of the algorithms.
- **Step 3: Scattering Sensitivity Test**
    - Exploring extremes: Investigate how sensitive the model is to variations in the scattering coefficient.

    - Comparative simulations: Perform simulations by first setting all scattering coefficients to 0% and then to 99%, keeping the absorption coefficients unchanged.
- **Step 4: Calibration of Reverberation through Absorption**
    - Global modification: The absorption coefficients are adjusted to bring the average differences in reverberation parameters between the simulated and measured results below 1 JND (Just Noticeable Difference, the minimum perceptible difference).
    - Priority to large surfaces: The first changes are applied to the materials covering the largest surfaces in the room, as even small variations generate noticeable effects.
- **Step 5: Calibration of Clarity through Scattering**
    - Use of sensitivity data: Based on the results obtained in Step 3, the scattering coefficients are adjusted.
    - Acoustic objective: These coefficients are adjusted to calibrate the average clarity parameters and bring them within 1 JND of the actual measured values.
- **Step 6: Local Refinement and Error Minimization**
    - Positional Analysis: The results are analyzed for each specific combination of position between the sound source and receiver.
    - Standard deviation optimization: The acoustic properties of specific key local surfaces are adjusted to minimize the standard deviation of the differences between the simulated and measured parameters.
    - Preservation of average values: These local corrections are made while ensuring that the global average values obtained in the previous steps are not altered.

In particular the second and third step is important in our case when the software is not a specialized room acoustics simulation software, which is the case for Blender VI-Suite. The present procedure facilitates the attainment of objective validation based on parameters, in addition to facilitating run-to-run verification. [42]

Figure 6.1 specifications for implementations conceptual map

6.3.3 Validation

In Katz's work, the term "validation" is primarily objective and internal to the method. The model is considered to be consistently calibrated with respect to the measured reference if the simulated parameters fall within 1 JND and the residual differences are within the run-to-run variability estimated in second step of the procedure. However, Katz is explicit on one point: complete perceptual validity also requires subjective listening tests, which are indicated in the paper as future work. Namely, to validate a VR acoustical mdoel, it is necessary the comparison between auralisations from the calibrated model and measured responses, and the possible role of the visual model in VR perception. It is evident that, in the absence of this crucial final stage, the model falls significantly short of achieving perceptual validation. [42]

6.3.4 Caveat and assumptions

The procedure was derived through the modelling and calibration of two case studies: an outdoor amphitheatre and the Saint-Germain-des-Prés church in Paris. While the amphitheater emphasised the significance of simulation variability, the intricate indoor environment of Saint-Germain-des-Prés underscored the necessity for a structured calibration workflow, culminating in the development of the six-step methodology. Indeed, the environments under consideration have been found to possess relatively well-distributed absorption and reverberation times, which appear to be largely independent of scattering. However, different environments (such as outdoors or spaces characterised by extremely uneven absorption), the procedure may not be able to transfer "1:1" without verification. Furthermore, the method presupposes that calibrating a discrete set of positions also provides confidence in unmeasured positions, and that the chosen parameters are sufficient to guarantee perceptual realism. [42]

7 Acoustic Simulation Workflow in Odeon and AuVi

7.1 Virtual model development

The construction of architectural spaces is facilitated by the utilisation of standard Blender modelling tools, wherein each surface is represented as a polygonal mesh. These mesh elements serve a dual function: they define the visual appearance of the space and simultaneously act as acoustic reflectors or absorbers, depending on their assigned material properties.

The development of simplified models that represent the surfaces without thickness that the sound generated inside the room encounters is essential. The following two models will be constructed: the first will incorporate the development of cylindrical diffusers and Schroeder's diffuser in three dimensions with its own thickness; the second will consist of flat surfaces that replace the two elements, but which will still describe their acoustic behaviour when the absorption and scattering coefficients are inserted into the material specifications.

The models to be prepared in Blender for importation into Odeon will comprise separate layers for each surface. Thereafter, the material properties for each of these will be specified in Odeon itself. The models to be utilised for the purpose of testing the AuVi - Suite in Blender will be a single mesh containing all the surfaces. For each surface, the material will be specified using the AuVi material section.

The model of the space must be manifold, especially for the acoustic simulation in Blender. Hence I had to work on the modeling part to obtain a manifold mesh until every edge belongs to exactly two faces, the mesh has no holes or gaps, has no internal or overlapping geometry, and has consistent face normals. The latter point, namely the normals orientation, because the AuVi simulation is based on a ray tracing technique, and only with faces pointing outward (not randomly flipped) the sound can be simulated, scattered, absorbed properly.

If these best practices are not followed, when programming in the geometry nodes part the acoustci simulation, the procedure will block and a system error saying that the room is not recognizable will appear.

In the picture showing the model that has been selected in the visibility of the room in *Object (Object properties) > Viewport Display > Display as >* to be displayed as *wire,* so it can be seen clearly the differences of the inside of the two models.

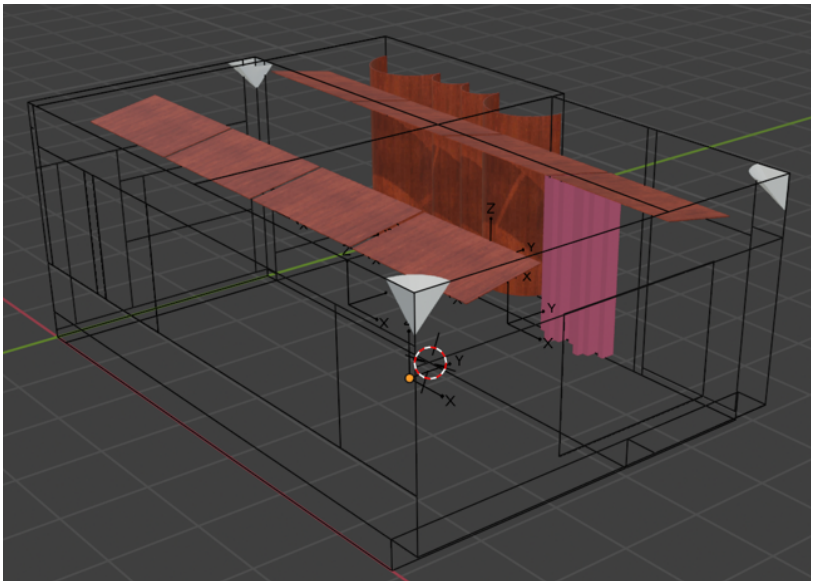


Figure 7.1 complex model, with diffuser components in 3D with 5,837 faces

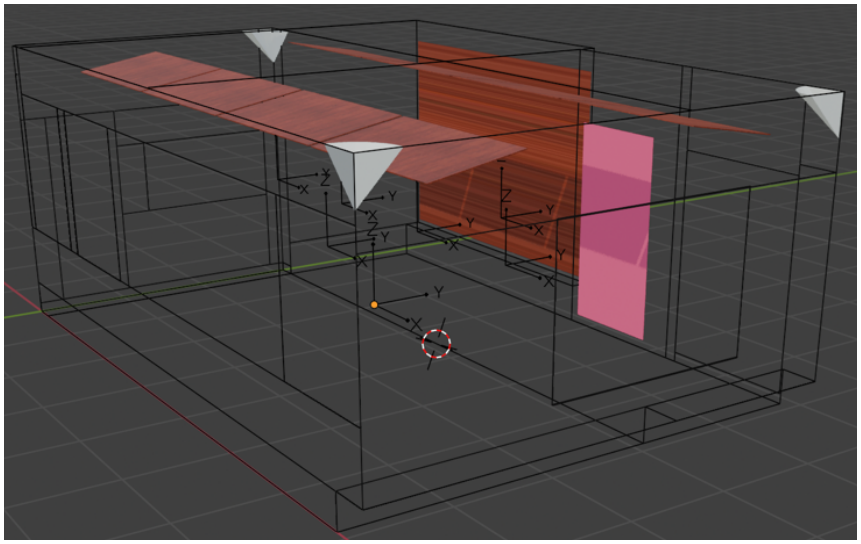


Figure 7.2 simplified model with 2D panels standing out for diffusing components with 147 faces

7.2 Room Material Mapping

Table 7.1 All materials present in the room with scattering coefficient averaged in the range of 500 Hz - 1 kHz

Component	Material	Scattering coefficient (s)	Absorption coefficient (α)						
			125 Hz	200 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
walls	Plaster. gypsum. or lime. smooth finish on lath (Harris, 1991)[43]	0.01	0.14	0.1	0.06	0.040	0.04	0.03	0.02
Stone walls	Brick. Unglazed (Harris, 1991)[43]	0.2	0.03	0.03	0.033	0.044	0.05	0.07	0.09
ceiling	Knauf Danoline Plaza. sound-absorbing forated plasterboard (manufacturer data)[44]	0.01	0.450	0.6	0.7	0.600	0.55	0.45	0.35
floor	Wood parquet in asphalt on concrete (Harris. 1991)[43]	0.01	0.042	0.042	0.074	0.063	0.063	0.074	0.084
windows	Double glazing. 2-3mm glass. >30mm gap (Fasold Winkler, 1976)[45]	0.01	0.15	0.05	0.03	0.04	0.02	0.02	0.02
studio window	Double glazing. 2-3mm glass. >30mm gap (Fasold Winkler, 1976)[45]	0.01	0.15	0.05	0.03	0.04	0.02	0.02	0.02
steel doors	Solid timber door (akustik.ua)[43,46]	0.01	0.140	0.100	0.06	0.08	0.1	0.1	0.1
bench	Wood parquet in asphalt on concrete (Harris. 1991) [34]	0.01	0.04	0.04	0.07	0.06	0.06	0.07	0.08
cylindrical diffusing elements	Thin Plywood paneling (Ref. Dalenback. CATT)	0.89 *	0.42	0.21	0.1	0.08	0.06	0.06	0.06
Ceiling suspended diffusive panel	8mm wooden ceiling suspended panel (Cox & D'Antonio, 2003)[47]	0.01	0.05	0.08	0.08	0.1	0.1	0.1	0.05
Schroeder diffuser	Thin Plywood paneling (Ref. Dalenback. CATT)	0.87 *	0.05	0.08	0.08	0.1	0.1	0.1	0.05
Bass traps	pressed polystyrene EPS (Cox & D'Antonio, 2004)[47]	0.01	0.01	0.01	0.02	0.02	0.035	0.06	0.07

The table presents the scattering coefficients of the Schroeder diffuser and the cylindrical panel, expressed as average values over the 500 Hz–1 kHz frequency range. A comprehensive description of their scattering and diffusive behaviour can be found in Appendix B.

7.3 Odeon

ODEON is a software program that simulates room acoustics. It is based on geometrical acoustics methods, primarily employing ray tracing to model sound propagation within enclosed spaces. [48] It facilitates the estimation of critical acoustic parameters, including reverberation time, clarity, and sound pressure distribution, by modelling the interaction of sound energy with room surfaces through reflection, absorption, and scattering.[39] ODEON is a widely utilised software in the fields of architectural acoustics and room design. It facilitates the evaluation of acoustic performance and the virtual testing of different materials and geometrical configurations prior to physical implementation.

The geometrical models used for the acoustic simulations were developed in SketchUp and subsequently exported in *.par* format for implementation in ODEON. SketchUp is a 3D modelling software widely employed for architectural and spatial design due to its intuitive interface and efficient handling of geometric constructions. In the context of room acoustics simulation, it allows the creation of accurate enclosure geometries that can be easily prepared for acoustic analysis. The use of SketchUp ensured a streamlined workflow from geometric modelling to acoustic simulation, enabling the generation of compatible *.par* files required for import into ODEON.

7.4 Set-up on Odeon

The first model that it has been used to calibrate all the parameters has been the simplified one with bidimensional diffusers panel. The configuration of the simulation space at Odeon was established through the consistent calibration of calculation parameters, the arrangement of the source-receiver layout, and the calibration of auralisation settings. Initially, the calculation parameters of the acoustic model were established (Figure 7.3). The length of the impulse response was set to 1000 milliseconds with the aim of including the entire energy decay of the reverberant field. The number of late rays was increased to 200,000, well above the recommended value, in order to ensure greater statistical stability of the model, especially at low frequencies. The permitted ray loss was maintained at 10%, thereby ensuring a balance between accuracy and calculation times.

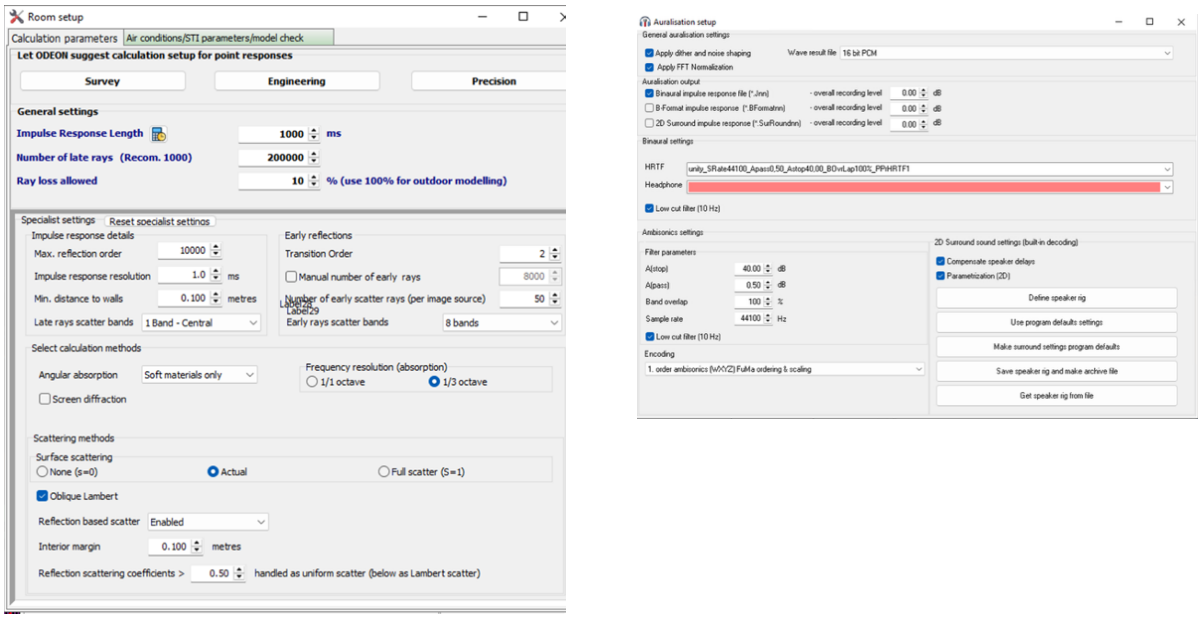


Figure 7.3 on the left it is shown the Room setup section for Odeon simulations, on the right Auralisation setup in Odeon showing the binaural impulse response configuration, selected HRTF, filtering parameters and ambisonic

In relation to the reflection modelling, a maximum order of 10,000 and a temporal resolution of the impulse response of 1 millisecond were utilised. The transition between the image method and ray tracing was managed with a transition order equal to 2, while early diffusion was described using 50 scattering rays for each image source distributed over 8 frequency bands. The analysis was conducted with a spectral resolution of one-third octave, thus ensuring the maintenance of adequate frequency detail without the imposition of excessive computational burden. From the perspective of diffusion, the real scattering mode of the surfaces was activated, thus avoiding the assumptions of both perfectly smooth surfaces and total scattering. The reflection-based scattering model was also enabled, and a Lambert-type oblique distribution was adopted to represent the diffusion of sound energy on internal surfaces in a more realistic manner.

Concurrently with defining the calculation parameters, the spatial distribution of the source and receivers was configured, as shown in Figure 7.4. The experiment utilised a point source positioned at ear level, with the receiving points arranged along the primary axis of the room and in lateral positions relative to it. All receivers were oriented towards the main source, thus enabling consistent evaluation of the sound field both in the direct direction and in off-center areas. The configuration for binaural auralisation was subsequently completed, encompassing the initiation of the calculation of binaural impulse responses. This calculation was facilitated through the utilisation of an HRTF (Head-Related Transfer Function) that had been selected from the software databse. The FFT normalization filters and low-cut filters were kept active in order to ensure stability and realism in signal reconstruction. The receivers were configured, in the apposite panel shown in Figure 7.5, as single-point response receivers within jobs in binaural mode, allowing each point to be associated with a specific impulse response and thus a possible perceptual

evaluation of the simulated environment. The integration of advanced calculation parameters, realistic scattering modelling, targeted distribution of measurement points, and binaural configuration has resulted in a robust framework for acoustic simulation of the Odeon space.

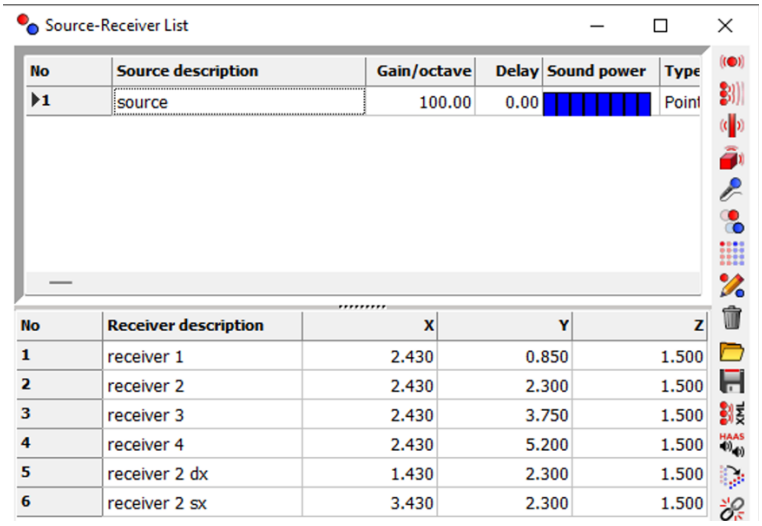


Figure 7.4 Source–receiver configuration in Odeon showing the position of the omnidirectional sound source and the spatial distribution of the receiver points used for the acoustic analysis within the room.

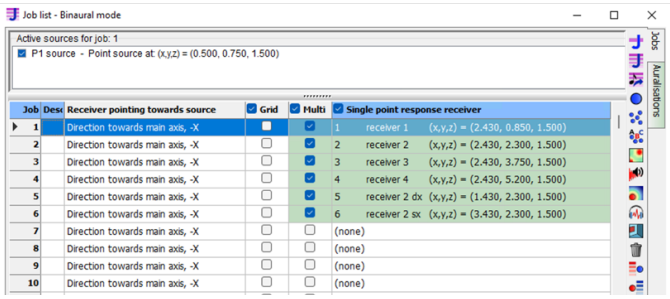


Figure 7.5 Job list in binaural mode showing the active point source and the configuration of single-point receivers oriented towards the main axis, used for binaural response calculations in the simulated environment.

7.5 Assigning materials to the model on Odeon

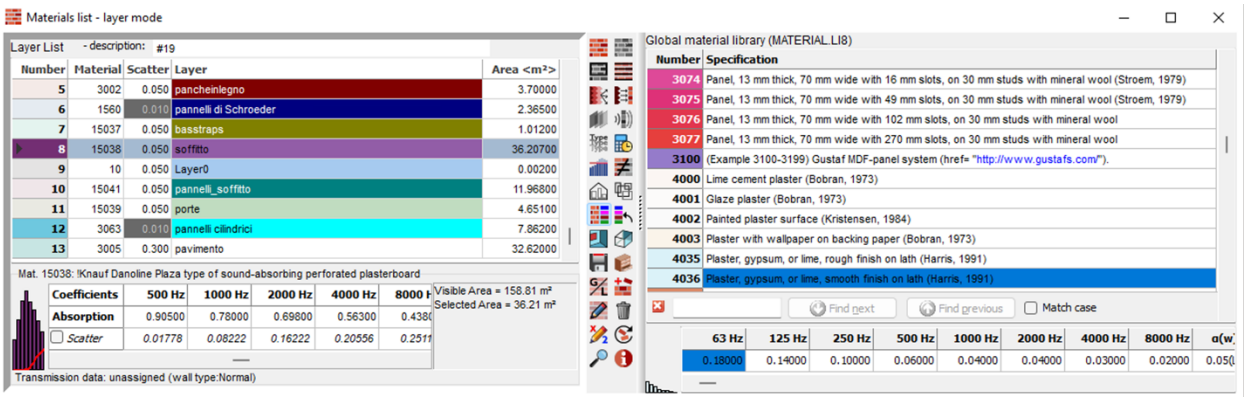


Figure 7.6 "Material list" section on Odeon 19 to assign material at each layer and to modify respctively all the scattering and absorption coefficients

In ODEON, the materials were assigned by associating each surface of the model with an acoustic material characterised by frequency-dependent absorption coefficients. These values were selected from the software's material library and technical literature, as depicted in Table 7.1, and were then adapted during the model calibration phase as in the table in appendix C which shows the coefficients after calibration. In the case of the complex model, the diffusing panels were modelled in three dimensions, so sound diffusion was generated directly from the geometry of the surfaces. In the simplified model, however, the same elements were represented as two-dimensional surfaces, and their diffusing behaviour was reproduced by assigning them appropriate, frequency-dependent scattering coefficients, which can be find in appendix B.

7.6 Calibration, simulation and data analysis

To this end, a series of iterative simulations were conducted with the objective of calibrating the numerical model until a difference of less than 1 JND was achieved with respect to the measured model. A comprehensive discussion of these findings can be found in Appendix C. The calibration process was conducted by adjusting the scattering coefficients of the materials present in the room, starting from the one reported in Table 7.1. These parameters were progressively adjusted through a systematic comparison between the T20 reverberation time curve returned by the simulated model and that obtained from the analysis of the experimental data of the real model. This methodological approach enabled a systematic reduction in the discrepancies between the simulation and the measurement, until a level of agreement was achieved that was perceptually indistinguishable according to the JND criterion. The calibration process was conducted utilising the simplified model, with the objective of deriving values of absorption for the calibrated model presented in Table 9.4. Consequently, these latter coefficients were designated as the definitive reference and were further employed in the simulation of the complex model, in which the diffusers were explicitly modelled in three dimensions.

In this manner, the configuration calibrated on the simplified model was transferred to the more detailed geometric model, ensuring consistency between the calibration phase and the final simulation phase.

7.7 Comparison of Parameter-Based and Geometry-Based Diffusion in Odeon

When diffusive elements such as QRD panels are represented through simplified two-dimensional surfaces, their geometrical structure is not explicitly included in the model. Consequently, the ray-tracing algorithm interprets these elements as flat reflective boundaries, inherently producing specular reflections. In order to compensate for the absence of geometrically induced diffusion, it is necessary to introduce frequency-dependent scattering coefficients. In this context, the phenomenon of scattering can be considered a substitute for the physical diffusion generated by the wells of the diffuser. The effectiveness of the diffuser is intrinsically frequency-dependent, and typically significant only above the design frequency range. [37,49]

Conversely, when the same diffusers are modelled in full three-dimensional detail, their morphology is explicitly resolved within the geometrical model. The resulting surface discontinuities naturally produce angular redistribution of reflected energy through geometric interactions with the incident rays. In this configuration, diffusion emerges directly from the geometry, rather than being imposed through material parameters. Consequently, the scattering coefficient no longer represents the primary diffusive mechanism and may be assigned a constant value, accounting only for residual effects such as small-scale roughness and construction tolerances not captured by geometrical acoustics.[37,49]

The *S_s* model, in which diffusion is introduced through the application of scattering coefficients to flat surfaces, reproduces the overall trend of the measured data, curve *M* in the legend , particularly at mid-to-high frequencies where the diffuser is expected to be effective. In contrast, the *S_c* model achieves a comparable agreement with measurements without reliance on frequency-dependent scattering as the principal diffusive mechanism. This finding indicates that, in scenarios where the diffuser geometry is explicitly delineated, diffusion emerges inherently from geometric interactions with the incident rays, rather than necessitating the imposition of diffusion through material parameters.

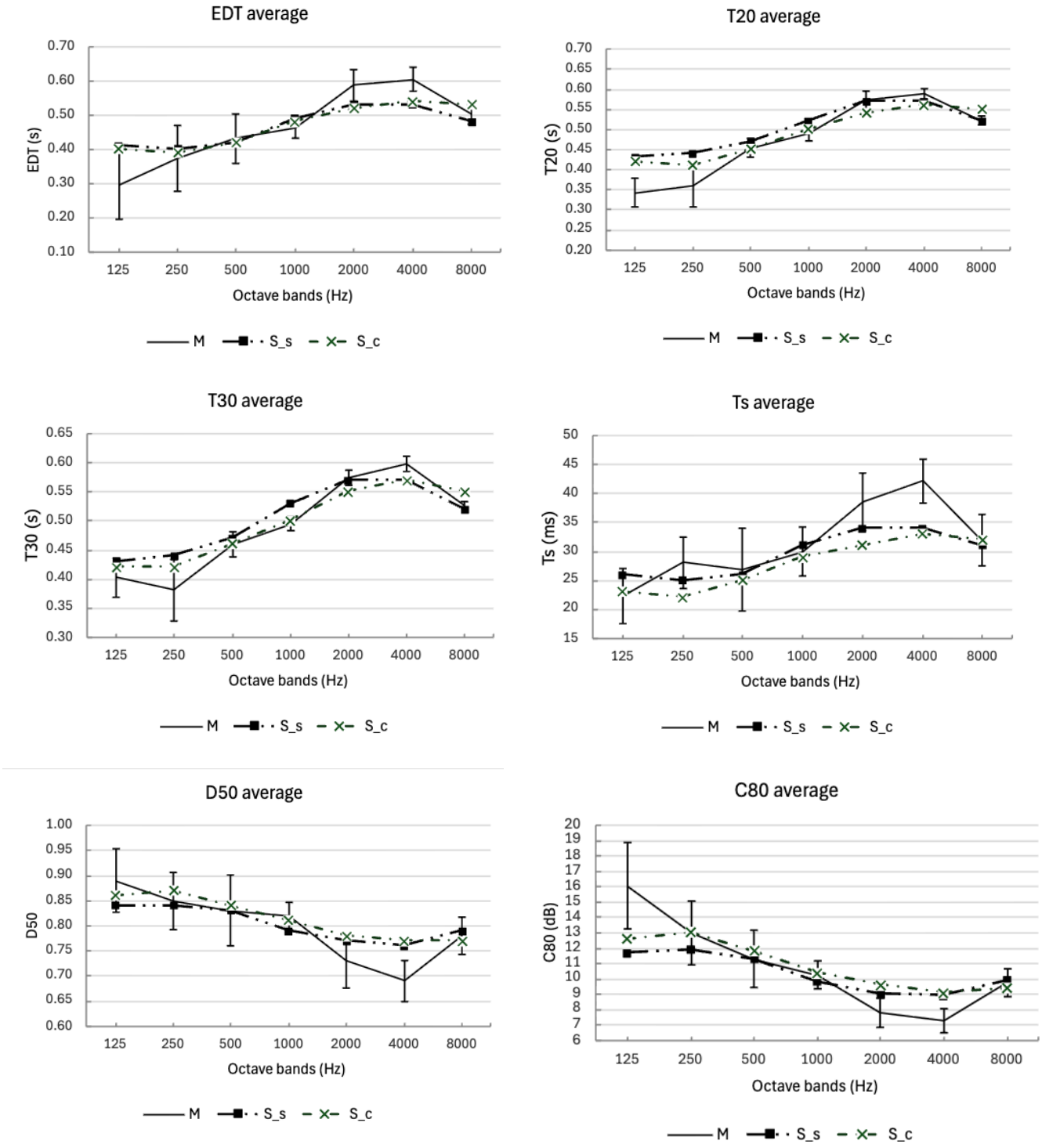


Figure 7.7 average comparison between the measured model with the complex and the simplified simulated models

Overall, the two simulations exhibit highly similar behaviour, accurately reproducing the general trend of the acoustic parameters observed in the measurements. In particular, the simulated and measured curves for the energy decay parameters (EDT, T_{30} and T_{20}) show a similar trend across the different octave bands. Reverberation times increase progressively from low to mid frequencies, reaching maximum values between 2 kHz and 4 kHz before decreasing slightly at around 8 kHz. From around 500 Hz onwards, the

differences between the simulations and measurements are generally small and often within the margin of error. The most obvious discrepancies are observed at low frequencies (125–250 Hz), where the simulations tend to overestimate decay times slightly compared to the measurements. This behaviour is typical of geometric acoustic models, which represent modal effects and actual system losses less accurately at lower frequencies. A similar trend emerges from the analysis of centre time (T_s). While the simulations accurately reproduce the growth of the parameter with frequency, the actual measurements reveal a more pronounced increase in T_s between 2 kHz and 4 kHz. This suggests the presence of more late energy than the simulations predicted.

This discrepancy is consistent with the findings regarding the clarity parameters, $D_{50}$and C_{80}. In both simulations, these parameters exhibit slightly higher values in the mid-high frequency bands, indicating that the energy is distributed over time in the early reflections. However, the actual measurements show a more pronounced decrease in D_{50} and C_{80} between 2 kHz and 4 kHz, indicating a greater late energy component in the actual space. Comparing the simplified (S_s) and complex (S_c) models shows that both produce very similar values for the main room acoustic parameters across most octave bands. This indicates that the frequency-dependent scattering coefficients employed in the simplified model can approximate the space's overall diffusive behaviour. Nevertheless, the complex model offers a more physically consistent depiction of sound diffusion, as the redistribution of reflected energy stems directly from the three-dimensional geometry of the diffusers rather than statistical scattering parameters.

7.7.1 Frequency plot analysis

In order to further validate this comparison, it was interesting to compare the IRs measured at SERMIG with the simulated IRs from the complex model and the simplified model using frequency plot analysis for the six receivers under consideration.

Overall, the simulations adequately replicate the prevailing trend of the frequency response in the mid and high frequencies, where the measured and simulated curves demonstrate significant overlap in both the overall slope and the distribution of spectral variations. This behaviour is consistent with the findings of the relevant literature regarding simulations based on geometric acoustics.

The most evident discrepancies between measurement and simulation, however, emerge at low frequencies, where the measured curves demonstrate more pronounced variations compared to the simulated ones. This phenomenon can be attributed to the presence of the room's natural modes, which dominate the sound field in the modal regime and are not accurately described by the ray-tracing models employed in acoustic simulation software. In the frequency range below the Schroeder frequency, the sound field cannot be considered diffuse and would require wave-based

modelling approaches to be reproduced with greater precision. [3]

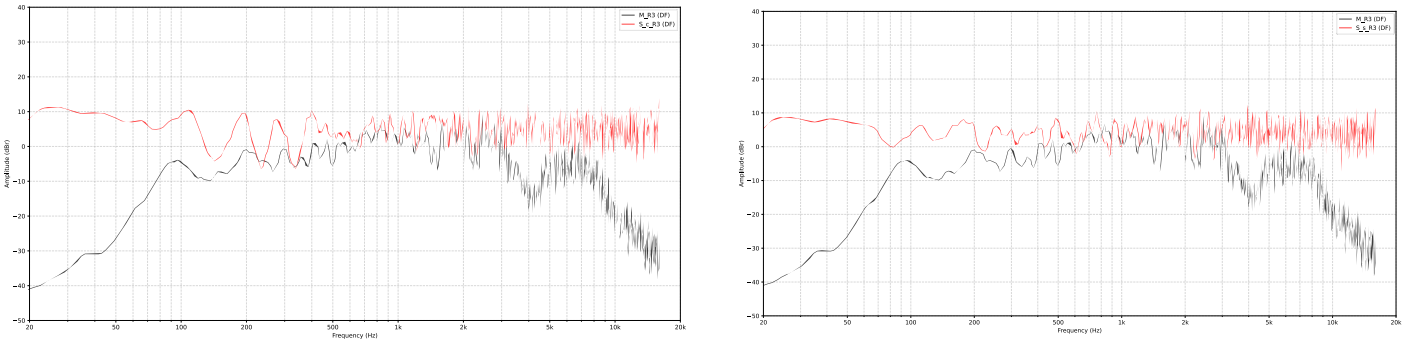


Figure 7.8 on the right the simplified model wih the respect to the measured one, while on the left the complex model compared to the measured

The comparison between all the IR shows that the simplified model provides frequency responses that are largely comparable to those obtained with the detailed geometrical model. The overall spectral behaviour remains consistent across the two modelling approaches, particularly in the mid- and high-frequency range. The main discrepancies with the measured responses are observed at low frequencies in both cases, confirming that they are primarily related to the modal behaviour of the room rather than to the modelling of the diffusive elements.

All the graphs are completely reported in Appendix O and Appendix P.

7.8 AuVi in Blender

The VI-Suite is a potent open-source add-on for Blender, a widely utilised 3D content creation platform. The software has been designed to function as both a pre- and post-processor for the Radiance and EnergyPlus programmes. The purpose of this design is to extend the capabilities of the Blender software, thus enabling the user to perform detailed environmental simulations and building performance analyses. The VI-Suite is distinguished by its characteristics of freedom, multi-platform functionality, and community-driven development, which aligns well with workflows that require flexibility, accessibility, and customisation.

As of version 0.6, the VI-Suite already offered a comprehensive set of tools for simulating, analysing and visualising environmental data directly within a 3D modelling context, encompassing for example the two-dimensional plotting of meteorological data from EnergyPlus weather file or dynamic sun path generation with real-time shadow casting and physically-based lighting simulations.
Whilst the VI-Suite version 0.6 did not yet incorporate acoustic simulations, significant advancements were introduced in version 0.7 with the development of AuVi, the acoustic analysis module. AuVi facilitates the simulation of reverberation times and impulse responses, with these simulations being based on the room geometry and material properties that have been defined within the Blender software[50].

This work was conducted using AuVi, the acoustic extension of the VI-suite version 0.7, which integrates the Pyroomacoustics Python library (version 0.9), provided on GitHub as a test build by Ryan Southall in February 2026.

Dr Ryan Southall is a Senior Lecturer in the School of Architecture & Design at the University of Brighton (UK). His academic work focuses on open-source tools for building environmental analysis, including lighting, thermal, ventilation, and acoustics, within Blender using tools like VI-Suite and AuVi.

Blender version 4.5.4 LTS is strongly recommended. The LTS (Long-Term Support) release ensures extended maintenance and improved stability. In contrast, non-LTS versions may introduce feature updates that can compromise compatibility and system robustness, particularly when interfacing with beta-stage add-ons such as the one adopted in this study.

7.9 Acoustical scenario reproduction in Blender

After the model is manifold and ready to continue the acoustic simulation in Blender we have to join all the faces of the rom, witout considering the bass traps, the diffusers and other furniture and furnishing. As you can see in the *scene collection* in figure 7.9, we have just one layer for all the room, its walls, ceiling and pavement, called “studio”, and the other layer divided sometimes in sub-collection are the accessories of the room.
To make visible the room from the AuVi nodes later, selecting the layer of the room, going deep in “Editor Type:properties” to see its Object properties, if the add-on properly it’s installed properly, the type of VI-Suite must be selected as an EnVi/AuVi Surface in the box shown in Figure 7.9.

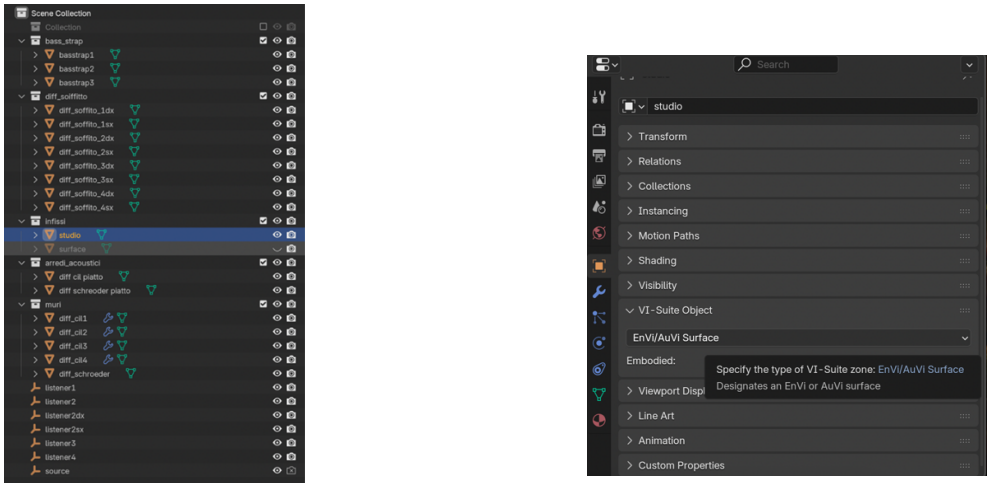


Figure 7.9 on the left Blender scene collection, on the right VI-Suite Object box for the object properties details

Each material of each surface has to be managed in the section *Material > Material proerties > VI – Suite Material* as shown in Figure 7.10. In this one after selecting *AuVi materia* in the *material type,* it is possible to specify the acoustic properties of the material, namely its absoprtion and scatter class. For both of these specification, namely the μ_s, the scattering coefficient, and the α, the absorption coefficient, it can be

chosen to select these material properties from the database of VI-Suite or to custom the coefficient as in Figure 7.11.
In this work some of the material properties were gently selected from the database resources other, in particular the scattering coefficient of the cilindrical panels and of the schroeder diffuser, having more measurements and material about it, custom properties were preferred.
It is even possibile in the same material to have the *AuVi absorption class* selected from the database and the *AuVi scatter class* as custom from the user, or viceversa.
It is worth examining this in depth to see the different possibilities that this choice can offer.

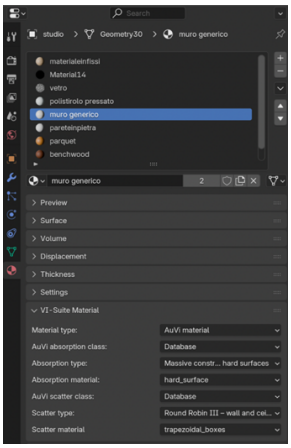


Figure 7.10 VI-Suite Material section

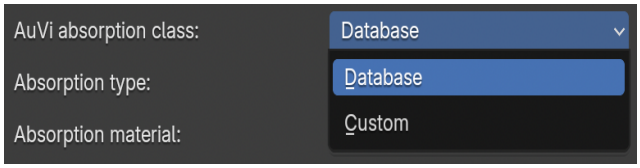
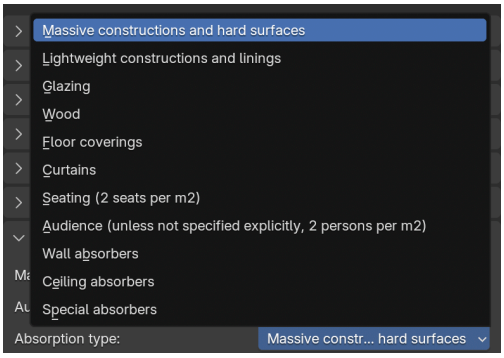


Figure 7.11 database or custom absorption and scattering class of the material

7.9.1 Database material

By selecting the database, we can define under the Absorption category how the material will behave in terms of sound absorption.
We have several of different choice (see Figure 7.12) , that seems from the forum that might be extended with future upgrades of the Add-on and its Python-library pyroomacoustics.
After specifying the material behaviour as an absorber, it can be selected, under *Absorption material,* the material name properly.



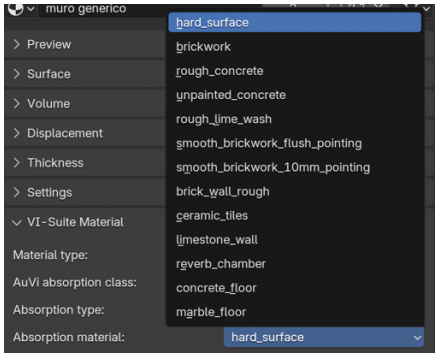


Figure 7.12 Absorption type option select window on the left and absorption material select window on the right

Regarding the *AuVi scatter class*, selecting the information present in the Database, under *Scatter type,* it can be selected one over these three options seen in Figure 7.13.
In particulare when the scatter type is the one of *Seating and audience* it means that it behaves as a component with an irregular shapes and soft surface created by chairs and people. While the *Round Robin III – wall and ceiling* refers to a standardized reference acoustic scattering model used to simulate how walls and ceilings diffuse sound in a room. So selecting the latter, the softwar applies predefined scattering coefficients that simulate the interaction of sound with architecturale surfaces like walls and ceilings.
As for the absorption, even for the scatter, under *Scatter material,* it can be selected *rect_prism_boxes* or *trapezoidal_boxes* as shown in Figure 7.14. The first option stand for a rectangular prism box behaviour, this is a six-faced 3D shape often reffer to as cuboid, which is often used as model to describe structures in scattering experiments and in physical simulations.

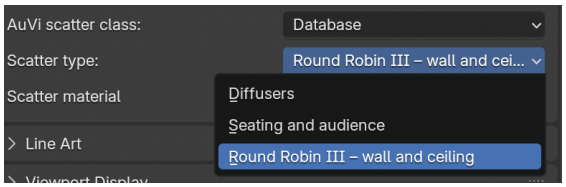


Figure 7.13 **Scatter type options from the database**

Figure 7.14 **Scatter material type**

7.9.2 Custom Material

In this work just two material are specified as *custom* since the scattering coefficients of the cilindrical diffusers and of the Schroeder diffuser has been previously studied more in details with AFMG Reflex.

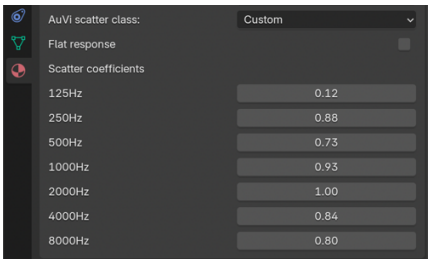


Figure 7.15 Scattering coefficient of the cilindrical diffuser specified in Blender

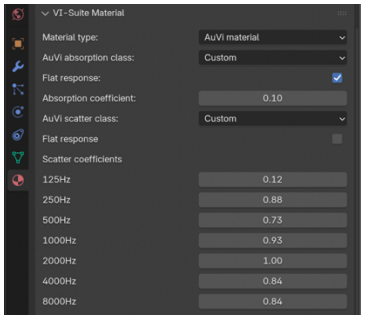


Figure 7.16 Scattering and absorption coefficients of the Schroeder diffuser specified in Blender

7.9.3 Listeners and source positioning

Getting deep in the simulation, the receivers, namely in our case the various microphone position, and the source must be collocated in the virtual space. In blender these can be represented es *Empties,* as you can see in Figure 7.17. In the context of the Blender software, 'empties' are distinguished as unique entities within a scene, despite their absence of geometric complexity. This signifies that they are devoid of surface, volume, or material. Consequently, they are not rendered, do not cast shadows, and are not visible in the final image. Notwithstanding this fact, they play a pivotal role in the workflow. The function of empties is to act as reference points in three-dimensional space, and they are utilised for the purpose of controlling or guiding other objects. These elements can function as targets for a camera or light, as alternative centres of rotation, or as support elements to create hierarchies between objects without the necessity of introducing visual elements into the scene. These tools are instrumental in facilitating organisation and control, a functionality that is particularly advantageous in the context of animations and the utilisation of constraints. In the viewport, these symbols are represented by simple geometric shapes, such as crosses, circles, and arrows. However, upon rendering, these symbols become completely invisible. In essence, these elements can be regarded as control "handles" in three-dimensional space, which facilitate the management of object behaviour without compromising the visual appearance of the scene

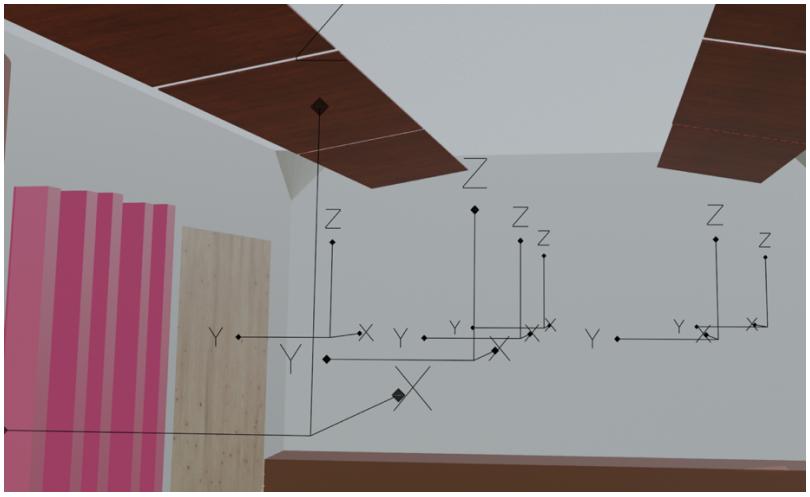


Figure 7.17 listeners and source positioning

Within the Blender interface, these objects can be added to the scene by selecting *Add* > *Empty* from the top-left menu. For the purposes of this work, the specific type of *Empty* is not critical; for instance, the *Plain Axes* can be used equivalently to *Arrows* or other *Empty* representations.

The only thing you mustn't forget before starting the simulation is to configure the *Object Properties* > *VI-Suite Object settings,* selecting the correct AuVi type for both the receivers and the source (which, in our case, is a single source): *Listener*, respectively (Figure 7.18) and *Source* (Figure 7.19).

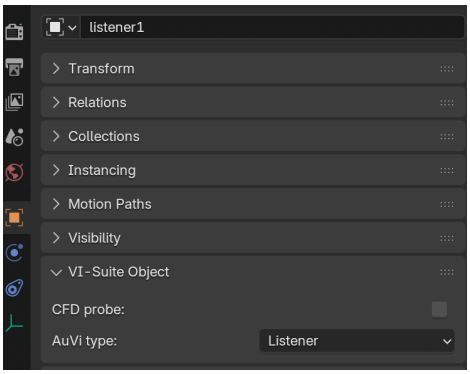


Figure 7.18 Selecting listeners as Listener type in the VI-Suite Object properties

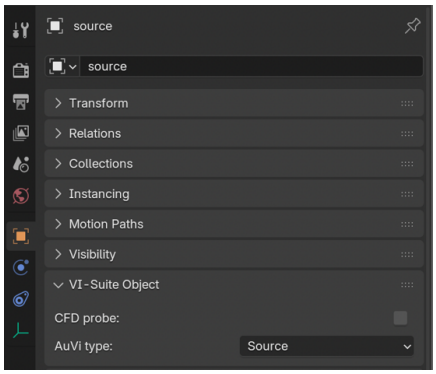


Figure 7.19 Selecting the source as a Source type in the VI-Suite Object properties

7.9.4 AuVi Geometry nodes programming

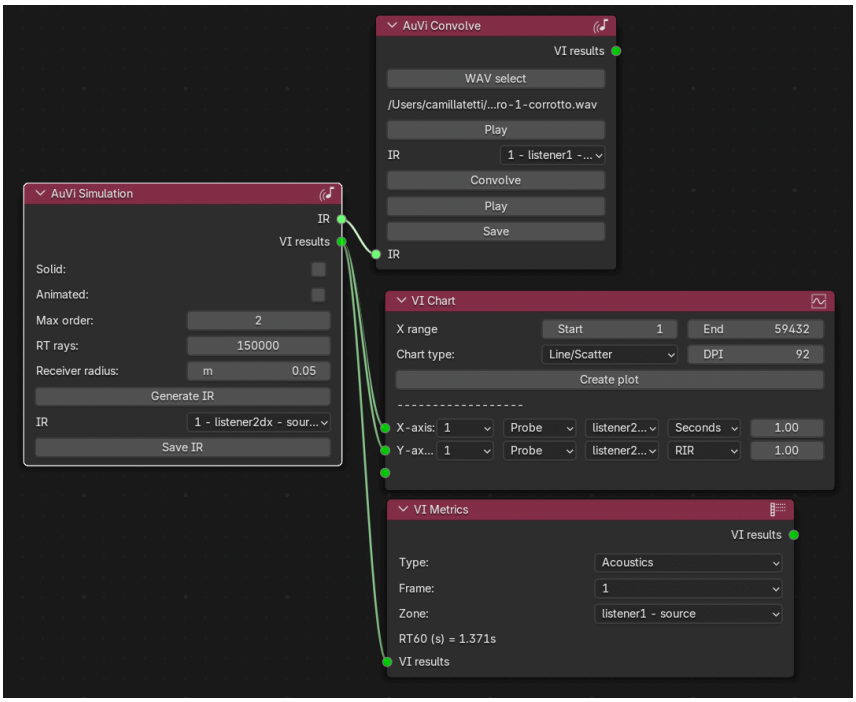


Figure 7.20AuVi geometry nodes for the acoustic simulation

Geometry Nodes in Blender were used to procedurally generate and control elements of the geometric model used in the AuVi acoustic simulations. This node-based workflow allows parametric modification of the scene while maintaining consistency across different simulation configurations.

This interface represents the node-based programming environment used to configure the simulation workflow. While the operations are visually organized through nodes, the underlying structure relies on Python code, where the acoustic calculations are performed using the Pyroomacoustics library, responsible for the simulation of sound propagation and the generation of impulse responses.

The AuVi Simulation node is the central component of the acoustic simulation workflow, as it is responsible for generating the impulse responses. From this node, there is the capacity to exercise control over several key parameters, including the Max Order, which defines the maximum number of reflections considered in the simulation. In this thesis, a value of 2 was adopted in order to maintain consistency with the simulation settings used in Odeon. The additional parameters include the RT rays, which determine the number of rays used in the ray-tracing algorithm of the acoustic simulation, and the Receiver Radius, which represents the minimum sensitivity radius of the receiver. Smaller values lead to a more precise estimation of the acoustic field. From the same node, it is also possible to select individual receivers. In addition, the corresponding impulse responses can be generated, and the results exported locally for further analysis.

The AuVi Convolve node facilitates the selection of a .wav audio file, which is then convolved with the impulse response generated by the simulation node. This process enables both the auditory evaluation of the simulated acoustics and the exportation of the resulting audio file. The VI-Chart node provides graphical visualisation of the impulse responses, thus enabling the analysis of their temporal decay as well as the generation of additional diagnostic plots that are useful for acoustic evaluation. Finally, the VI Metrics node displays instant estimations of average reverberation time values, providing a rapid overview of the simulated acoustic behaviour.

7.10 Calibration of the room with AuVi

The acoustic computation itself is carried out using Pyroomacoustics, an open-source Python library tailored for indoor acoustic signal processing. AuVi employs a hybrid simulation model, combining the Image Source Method (ISM) to accurately resolve early reflections with a ray tracing technique for simulating late reverberation. This dual approach facilitates a more comprehensive representation of both deterministic and stochastic aspects of sound propagation. The simulation parameters, including the number of image reflections, the quantity of rays cast, and the receiver detection radius, are fully configurable by the user.
It has been used the same numerical configuration used in the calibrated Odeon model to have coherence between the simulations.

7.11 Data analysis of the first simulation on Blender

Three impulse responses were generated for each receiver and analysed using the same MATLAB script that was used to process the objective measurements. As can be seen in Figures 0.27–0.32, the results are significantly higher than the actual measurements. However, the curves for the complex and simplified models are almost superimposed, suggesting that the discrepancy depends on the process by which AuVi generates or exports the impulse responses rather than on the modelling method.
This suggests a potential limitation of the simulation framework or the IR export process, which is likely optimised for internal auralisation within the software, but may be less reliable for objective analyses that

comply with ISO standards. Even when listened to, the decay of the simulated impulse responses appears unnatural, characterised by a strongly electronic pattern. Subsequent simulations were obtained by modelling the materials in Blender using the same absorption and scattering coefficients as the Odeon-calibrated model. To maintain this level of detail, the IRs had to be exported at 16 kHz, which is the add-on's default sampling frequency. While this limitation may introduce uncertainties at high frequencies, it does not explain the significant discrepancy observed in comparison to the experimental measurements, as indicated by the negative clarity parameter values.

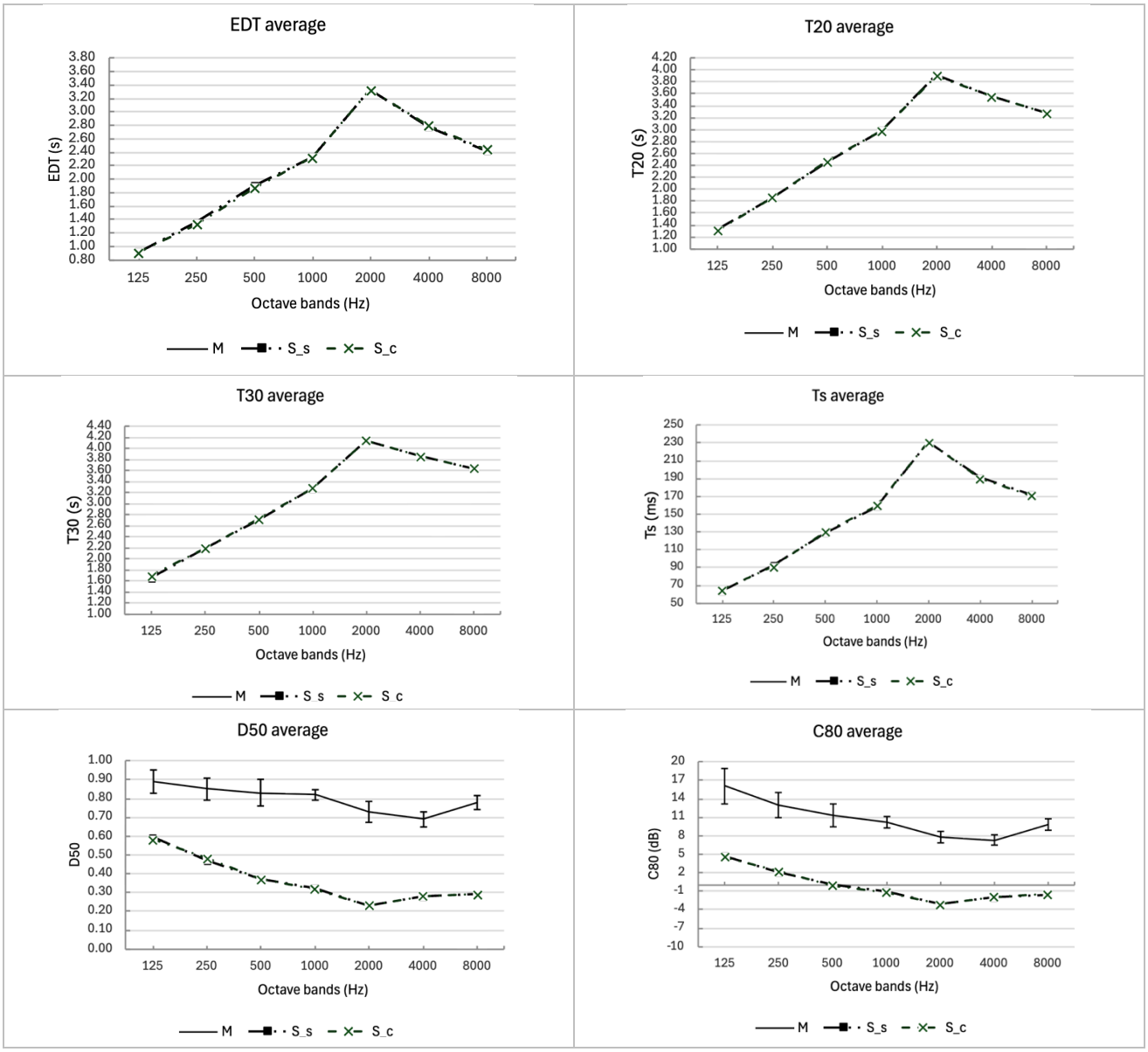


Figure 7.21 Key acoustic parameter plots obtained from the values with MatLab analysis for the first simulation om Blender

7.12 Second simulation on Blender

In this second attempt, rather than pursuing a simulation as close as possible to the model developed in Odeon, the decision was taken to build a configuration more compatible with the functionality and operational limitations of the AuVi add-on. In order to achieve this objective, the utilisation of materials present within the software database was employed, with the maintenance of flat scattering coefficients being consistent with the approach adopted in previous simulations within Odeon. It has been lowered the number of ray-tracing rays from 200k to 150k to lower the computational complexity. This choice entailed the abandonment of the simplified model with two-dimensional diffusers, which necessitates the definition of scattering coefficients contingent on the frequency band. The present attempt was therefore conducted exclusively on the complex model, in which the diffusers are geometrically modelled in three dimensions and the scattering coefficient is defined as flat, i.e. constant for all frequency bands. This approach is consistent with the observations that emerged from previous simulations, which did not reveal any particular sensitivity of the software to the distinction between geometric scattering and parametric scattering.
The absorption classes selected from the software database and the customised scattering values adopted are shown in Table 7.2.

Table 7.2 Absorption and scattering setting for AuVi simulation

Component	Material on Odeon	Absorption type	Absorption material	Scattering (flat)
walls	Plaster. gypsum. or lime. smooth finish on lath (Harris, 1991)[43]	massive constructions and hard surface	smooth_bruckwork_10mm_pointing	0.3
Stone walls	Brick. Unglazed (Harris, 1991)[43]	Massive constructions and hard surface	rrick_wall_rough	0.3
ceiling	Knauf Danoline Plaza. sound-absorbing forated plasterboard (manufacturer data)[44]	ceilig absorbers	ceiling_perforated_gypsium_plasterboard	0.05
floor	Wood parquet in asphalt on concrete (Harris. 1991)[43]	Wood	stage_floor	0.3
windows	Double glazing. 2-3mm glass. >30mm gap (Fasold Winkler, 1976)[45]	Glazing	double_glazing_inside	0.05
studio window	Double glazing. 2-3mm glass. >30mm gap (Fasold Winkler, 1976)[45]	Glazing	double_glazing_inside	0.05
steel doors	Solid timber door (akustik.ua)[43,46]	Massive constructions and hard surface	hard_surface	0.05
bench	Wood parquet in asphalt on concrete (Harris. 1991) [34]	wood	Wood_1.6cm	0.05
cylindrical diffusing elements	Thin Plywood paneling (Ref. Dalenback. CATT)	wood	plywood_thin	0.05
Ceiling suspended diffusive panel	8mm wooden ceiling suspended panel (Cox & D'Antonio, 2003)[47]	wood	Wood_16mm	0.05
Schroeder diffuser	Thin Plywood paneling (Ref. Dalenback. CATT)	wood	plywood_thin	0.05
Bass traps	pressed polystyrene EPS (Cox & D'Antonio, 2004)[47]	ceiling absorbers	ceiling_melamine_foam	0.05

7.13 Data analysis of the second simulation with AuVi

As might be expected, the add-on demonstrates a superior management of materials in its database when compared to the data shown in the previous simulation graphs. The data obtained thus far allows for a comparison of the frequency response curves of the measured model and the complex model simulated in Odeon. An analysis of the graphs reveals that AuVi significantly overestimates the decay at low frequencies. This results in elevated clarity and definition values, as demonstrated in Figure 7.22. The decay parameters demonstrate elevated values at low frequencies and exhibit a substantial decrease from 500 Hz onwards. Conversely, C_{80} is elevated, D_{50} and approaches unity, indicating that the energy arrives prematurely. [3] The uniformity of this behaviour indicates that the primary source of discrepancy cannot be ascribed to geometric modelling, materials, or their coefficients. Rather, it is attributable to the process of generating impulse responses within the simulation framework. This phenomenon can be attributed to the simulation method employed by Pyroomacoustics, which is primarily based on the Image Source Method. The Image Source Method is primarily concerned with the representation of specular reflections and may be less effective in modelling the diffuse field. [51] This finding serves to highlight the limitations of the simulation model, as evidenced by the behaviour observed at low frequencies. Specifically, AuVi simulations demonstrate elevated decay values at 125 Hz, followed by a precipitous decline commencing at 250 Hz. Such an abrupt trend is not typically observed in real measurements, where the variation in reverberation times at low frequencies is generally more gradual. This phenomenon could be attributed to the inherent limitations of geometric acoustics models in the modal regime, which characterises the sound field at low frequencies.[3]

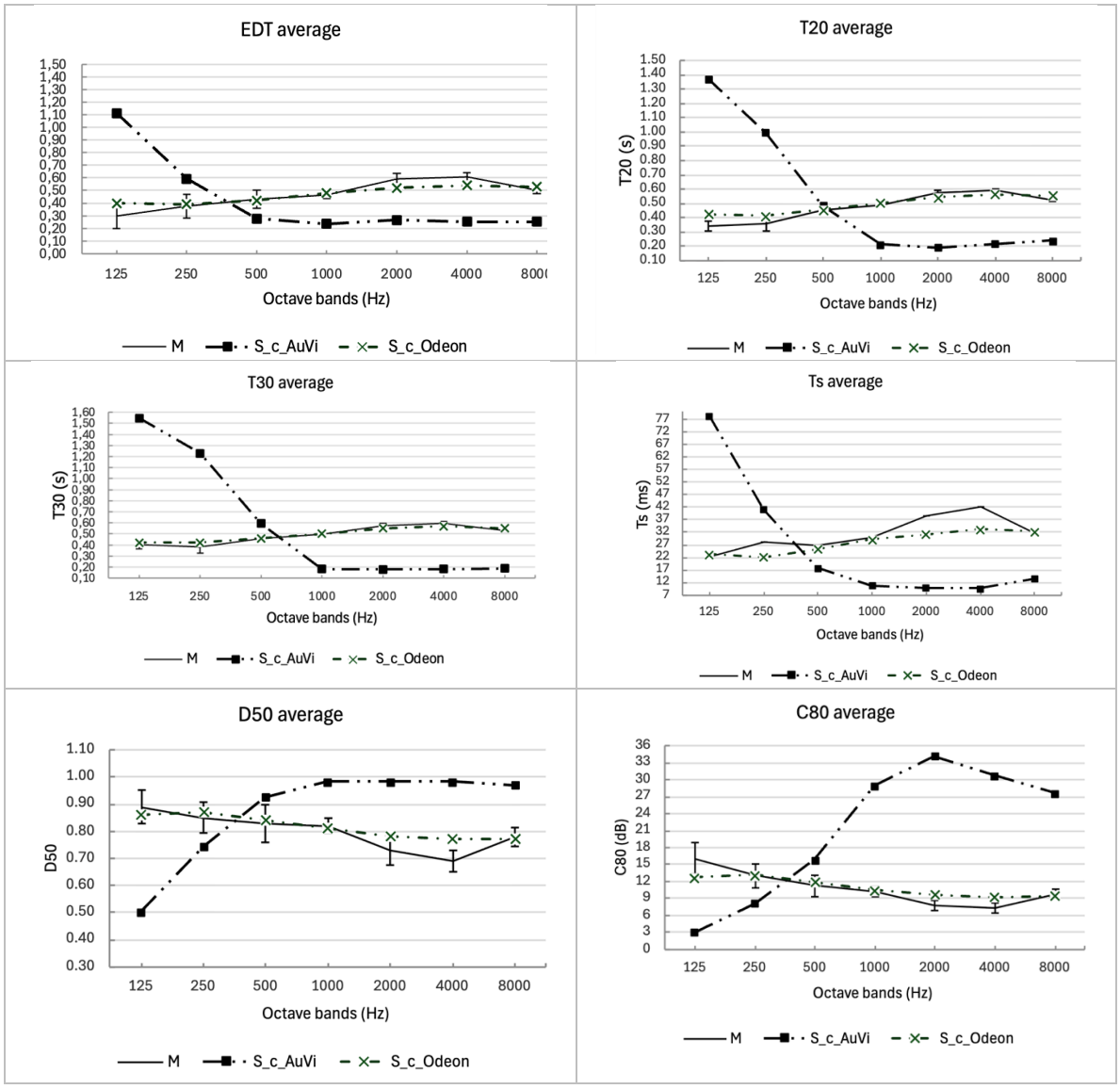


Figure 7.22 Key acoustic parameters value for the second simulation with AuVi

7.14 Comparison between Odeon and AuVi

Table 7.3 Synthetic comparison between the two software used for this research: Odeon and AuVi Add-on from VI-Suite in Blender

Feature	Odeon	AuVi-Blender (VI-Suite)
Software type	Professional software dedicated to architectural acoustic simulation	Energetic and acoustic simulation-oriented Blender open-source add-on
Simulation mathematical	Ray-tracing combined with image source method, optimized for room aocustics	Pyroomacoustics Python package, based on image source method
Geometric modelling	Simplified modeling features or CAD import	Complete 3D modelling thanks to Blender functionalities
Measurable acoustical parameters	EDT, RT, G, STI, T_{20}, T_{30}, D_{50}, C_{80}, C_{50} and further more	$RT_{60,\mathrm{mid}}$ and STI (but just for the latest beta version uploaded and not tested in this thesis)
Auralization	Integrated with different advanced convolution tools and reproduction set-up	Just possible through the *AuVi convolve* node which makes an IR based convolution
Graphic analysis	Dedicated advanced instruments	Some graphic analysis featurs , i.e. aplitude over time decay of the IR, throughout the *VI-Chart*
Simulation accuracy	High accuracy, scientifically validated	Experimental beta version
Accessibility	Commercial software	Free and Open-source software
Main scope of use	Architectural professional acoustic projects design	Research, experimental purpose and integration with 3D modelling

The synthetic comparison made in Table 7.3 has been constructed with the enpirical investigation done for this thesis and with the information obtained from the manual of this two software. [52,53]

8 Further simulations on Odeon

8.1 Positional dependencies of different configurations

The reliability of geometrical acoustics simulations depends heavily on the modelling assumptions adopted for interior elements and spatial boundaries. In complex environments, decisions regarding the extent of the model, the level of geometric detail and the representation of furnishings may significantly affect the predicted room acoustic parameters. Sluyts et al. demonstrate that, in some cases, truncating portions of a large geometrical model can have a negligible influence on local acoustic parameters. However, in other cases, particularly when the distribution of absorption changes, measurable deviations occur. Their study shows that the spatial distribution of absorbing surfaces and the geometric configuration of the model directly affect simulated reverberation behaviour. [54] This underlines the importance of systematically assessing the influence of modelling simplifications on acoustic predictions. Similarly, Nguyen shows that furniture cannot be treated as acoustically negligible.

Its absorbing and scattering effects can significantly alter reverberation time, speech clarity, and sound strength. [55]Notably, the study reveals that the acoustic contribution of furniture depends on the absorption properties of surrounding surfaces, especially suspended ceilings. This interaction effect indicates that interior elements must be evaluated not only individually, but also in relation to the room's overall absorption balance. Following a comparison of simulations conducted in Blender and ODEON considering both modelling strategies, three-dimensional diffusers with intrinsic geometric scattering and two-dimensional diffusers with parametrically assigned scattering, the analysis was further developed using ODEON as the reference environment. This was due to the software's greater sensitivity to variations in absorption and scattering, as well as its established validation in geometric acoustics simulations. In line with the findings of Sluyts and Nguyen, a systematic sensitivity analysis of the individual internal components of the room was deemed necessary. The investigation was conducted using two models: the complex model (S_c), which featured 3D-modelled diffusing panels, and the simplified model (S_s), which featured two-dimensional diffusers. The following configurations were analysed progressively: an empty room (S_{empty}), a room with bass traps only (S_{bass}), a room with suspended reflective panels only (S_{refl}) and a room with diffusing elements only (S_{diff}), including cylindrical panels and Schroeder diffusers. The latter configuration was studied in three and two dimensions.

Finally, all configurations were compared with both the complex and simplified models to evaluate the effect of each element and the influence of the modelling strategy on the resulting acoustic parameters.

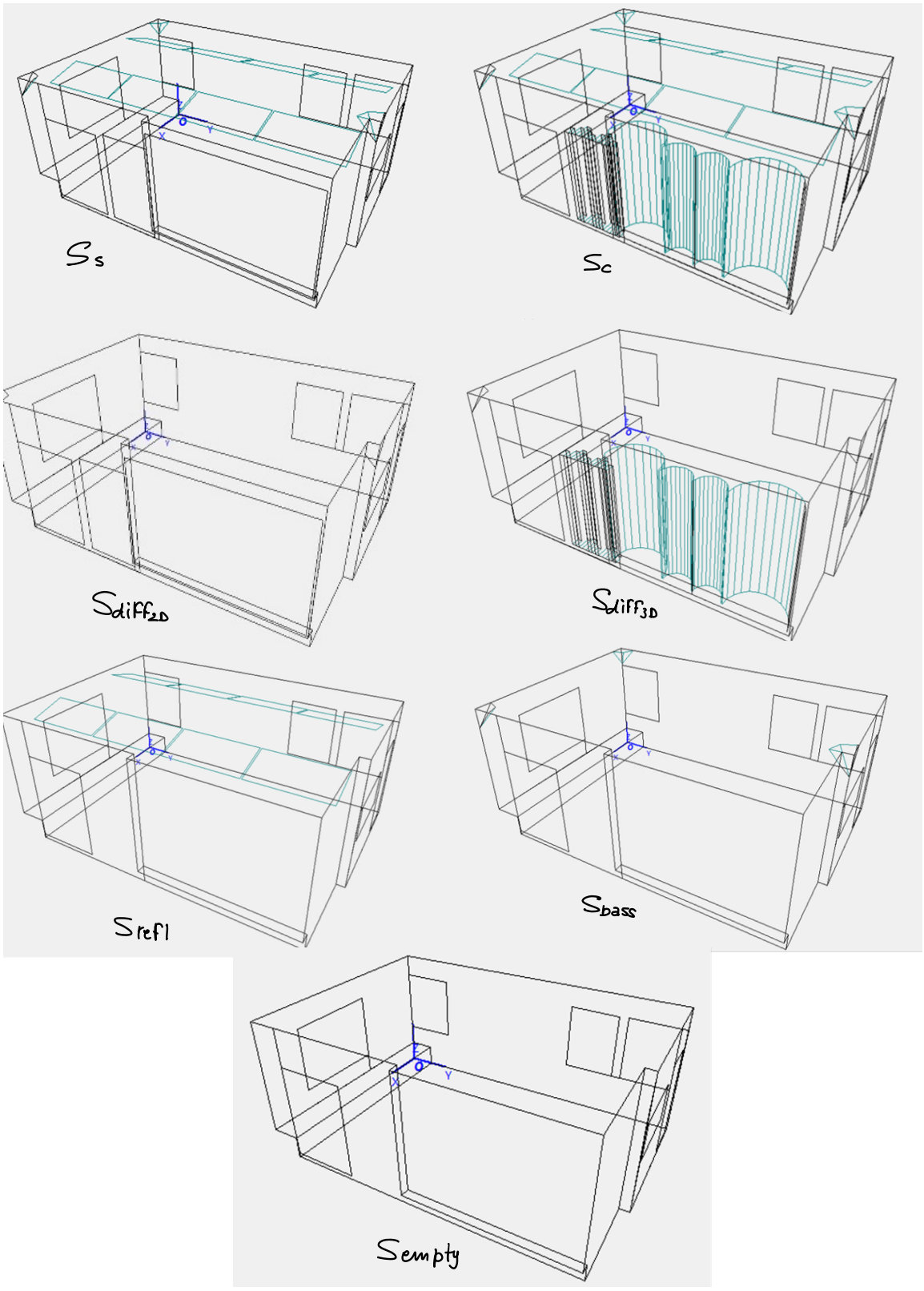


Figure 8.1 All different configurations of the room simulated on Odeon

8.2 Different configurations of the complex model of the room

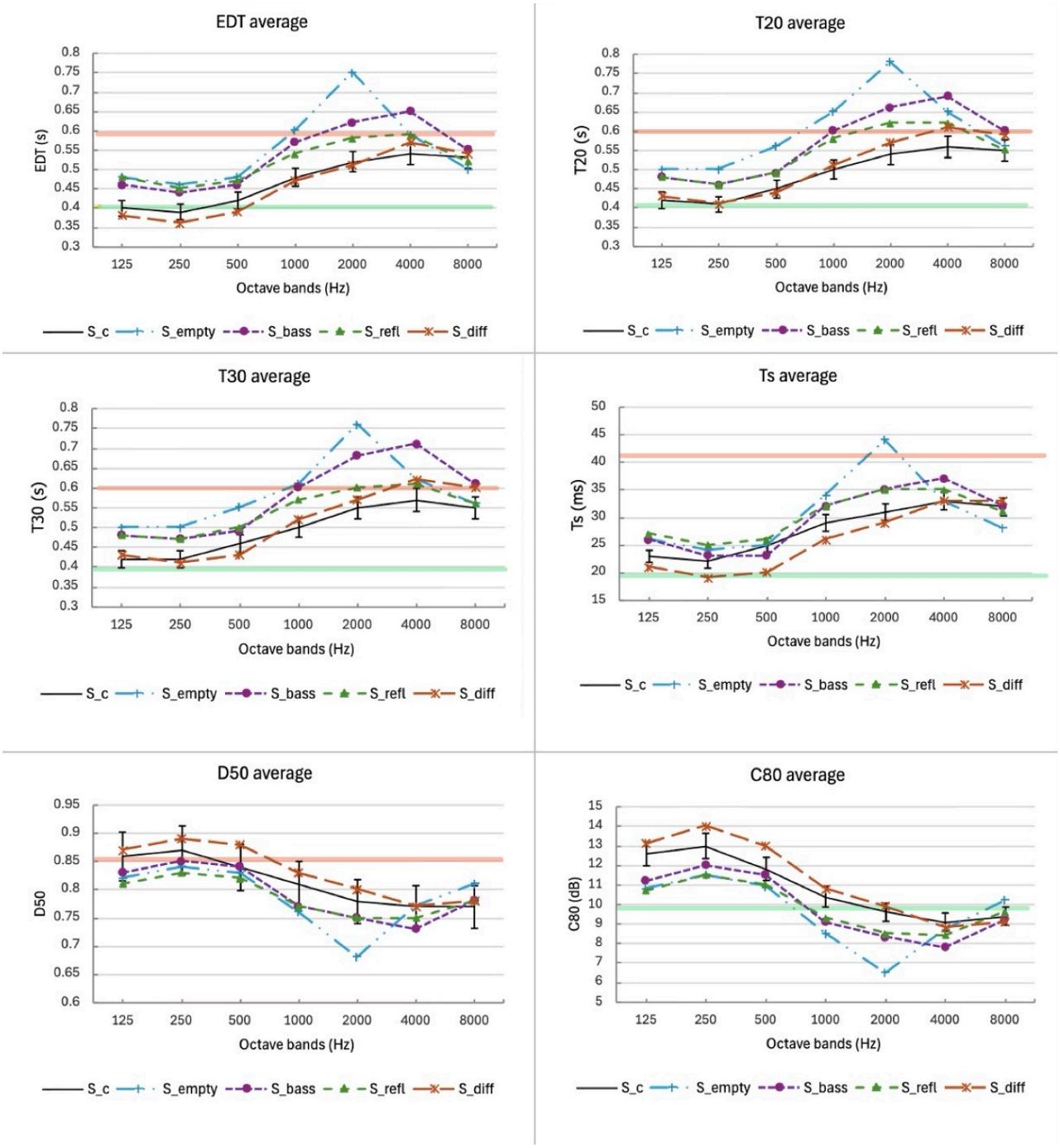


Figure 8.2 comparison of the key acoustic paramters for the different configurations with the respect to the complex model. Optimal parameter ranges are defined by a maximum (red) and minimum (green), based on literature and measured model analysis.

A comparison of the various configurations with the complex model (S_c) reveals how each introduced element contributes differently to the sound field, with effects that vary significantly with frequency. The empty room configuration (S_{empty}) is the most reverberant and least controlled case by definition, with a strong positive imbalance in the decay parameters. This leads to a negative imbalance in the clarity parameters.

Looking at the graphs, it is interesting to note that the difference between configurations is more pronounced in the 1000–4000 Hz band. This suggests that the mid-high behaviour of the room's surfaces most affects acoustic performance, while the difference decreases at low frequencies. Within this band, the 2000 Hz range is certainly the most sensitive.
As expected, the empty room deviates from the complete furnished model, whereas the room containing only cylindrical and Schroeder panels coincides closely with the furnished model.

$$\Delta P_i(2\ kHz) = P_i(2\ kHz) - P_{S_c}(2\ kHz) \tag{8.1}$$

Making clear this concept, we can define a delta to consider as in the equation 8.1, to see for 2 kHz, since it is the most sensitive frequency in this analysis, that for the empty room $\Delta C_{80,empty} \approx -3.3\ dB$, while for the one with the diffusive panel $\Delta C_{80,diff} \approx -0.1\ dB$.

We can see that the reduction in clarity index of over 3 dB in the empty room is acoustically significant, as it exceeds the JND of 1 dB compared to the configuration with diffusers. Conversely, the presence of speakers reduces the reverberation time gap by almost 80% compared to an empty room. So for the empty room $\Delta T_{20,empty} \approx +0.26\ s$ and for the model with just the diffusive panel is $\Delta T_{20,diff} \approx +0.05\ s$. We can say that the trend represented by the inequality 8.2 is evident, in order of influence from the most critical to the closest to the complex model.

$$S_{\text{empty}} > S_{\text{bass}} > S_{\text{refl}} > S_{\text{diff}} \approx S_{\text{c}} \tag{8.2}$$

The selective introduction of bass traps and reflective panels only partially reduces these deviations. In contrast, the configuration with diffusing elements (S_diff) most closely resembles the behaviour of the complex model, with Δ values close to zero for T_s and EDT, and minimal deviations for C_{80}. The optimal ranges for each parameter are indicated by a maximum value, shown in red, and a minimum value, shown in green. These values are based on a review of the literature and an analysis of the values obtained from the measured model.

A receiver-by-receiver analysis is provided in the appendix.

8.3 Different configuration of the simplified model of the room

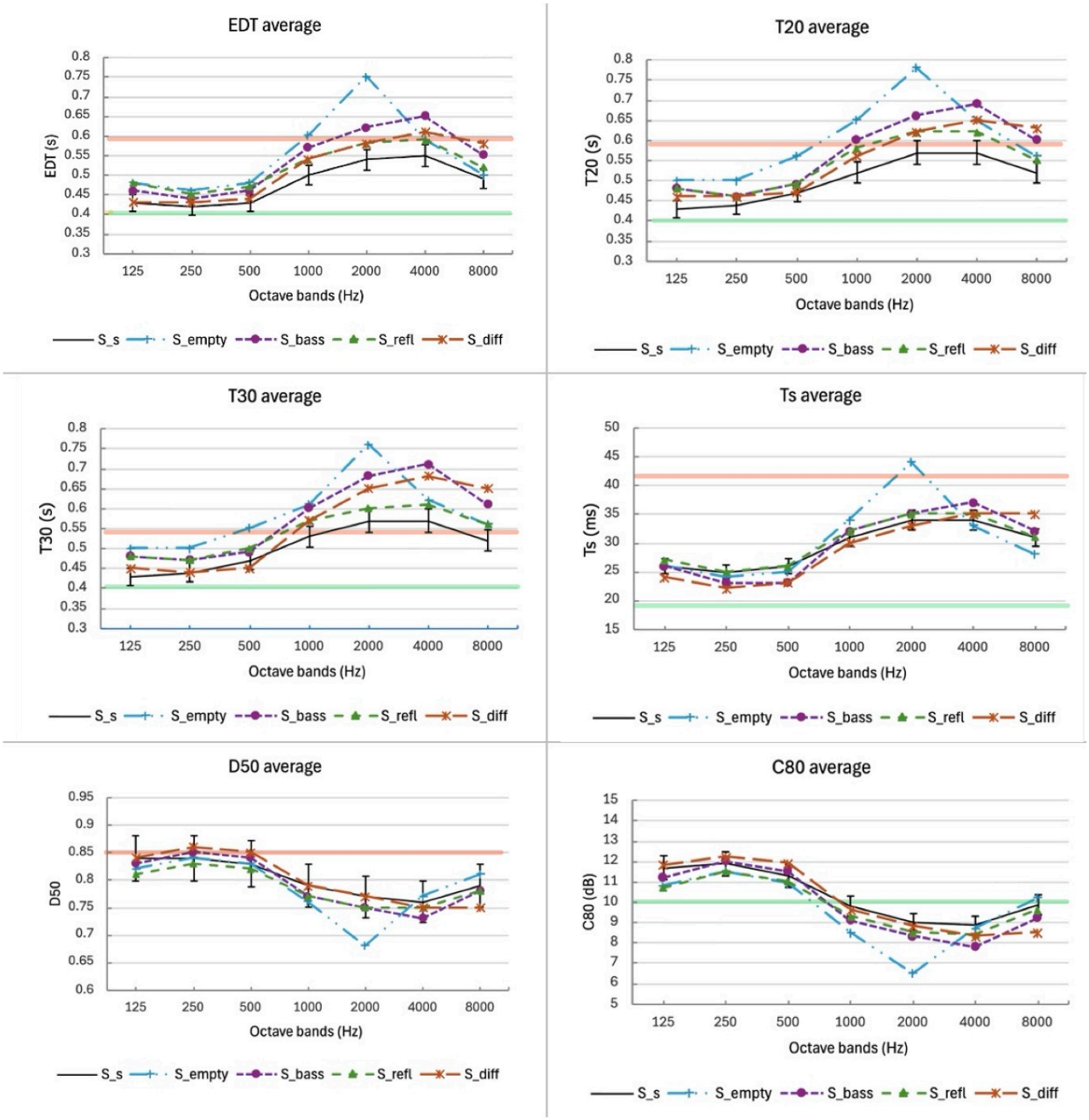


Figure 8.3 comparison of the key acoustic paramters for the different configurations with the respect to the simplified model. Optimal parameter ranges are defined by a maximum (red) and minimum (green), based on literature and measured model analysis.

As expected, the analysis of the complex model with three-dimensional diffusers shows that the general trend of the different configurations compared to the complete configuration is qualitatively similar, even in the simplified model.

The differences between the configurations are therefore greatest in the 1000–4000 Hz band, with particular sensitivity around 2000 Hz. The empty room configuration is the one that differs most from the complete model, while the configuration with only diffusing panels differs least.

However, the latter case is no longer acoustically comparable, and differences emerge from comparing the two references that should be highlighted. Unlike in the comparison with the complex model, the configuration with only diffusers is no longer neutral compared to the reference. This is evident from the deltas on the C80 parameter, which were previously almost zero.

$$\Delta P_i(2\ kHz) = P_i(2\ kHz) - P_{S_s}(2\ kHz) \quad (8.3)$$

Defining again our delta as in equation 8.3, since 2 kHz is the most sensitive frequency in this analysis, $\Delta C_{80,diff} \approx -0.7\ dB$.
In addition to analysing the general trend and net deviation in terms of clarity compared to the 2D model, it can be seen that the same configuration shows more pronounced deviations, particularly in terms of decay parameters.
In fact, observing the reverberation times of the various configurations shows that even the energy parameters no longer coincide with the reference for the configuration with only 2D diffusers, showin a delta of $\Delta T_{20,diff} \approx +0.08\ s$.

A comparison of the two references reveals that three-dimensional modelling of diffusers introduces an energy contribution that is not fully equivalent to that obtained through the parametric assignment of scattering. While the 2D model satisfactorily reproduces the clarity parameters, differences in decay times suggest that the spatial interaction between incident waves and diffusing morphology is underestimated when it is not explicitly modelled.

The optimal ranges for each parameter are indicated by a maximum value, shown in red, and a minimum value, shown in green. These values are based on a review of the literature and an analysis of the values obtained from the measured model.
A receiver-by-receiver analysis is provided in the appendix.

8.4 Frequency plot analysis of the different simulated IRs

As is evident in the tobjective acoustical parameter analysis, a significant proportion of these observations are corroborated when comparing the impulse responses obtained by simulating the various setups in Odeon.
Namely, observing Figure 8.4 and Figure 8.5, when comparing the impulse responses, the room equipped with three-dimensional diffusers demonstrates behaviour that is almost identical to that of the fully

modelled three-dimensional room. However, when contrasting the behaviour of two-dimensional diffusers with that of the simplified model, a significant discrepancy emerges.
The lower magnitude of S_s compared to both S_{diff2D} and S_c can be attributed to the simplified 2D representation of the diffusive panels within the complete room model, which leads to an underestimation of scattering effects and a reduced amount of reflected energy reaching the receiver.

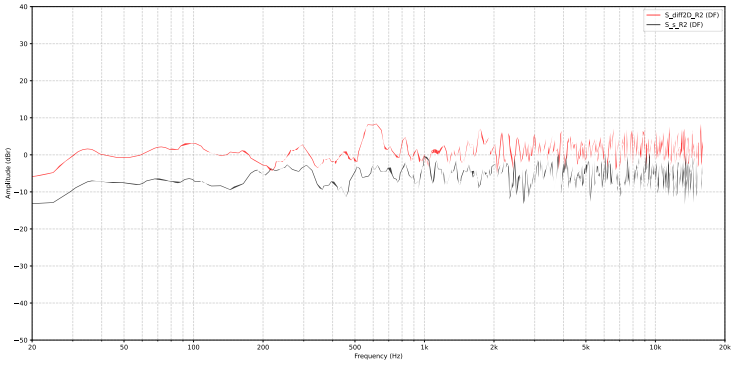


Figure 8.4 Frequency response plot comparison between S_{diff2D} and S_s from receiver 2

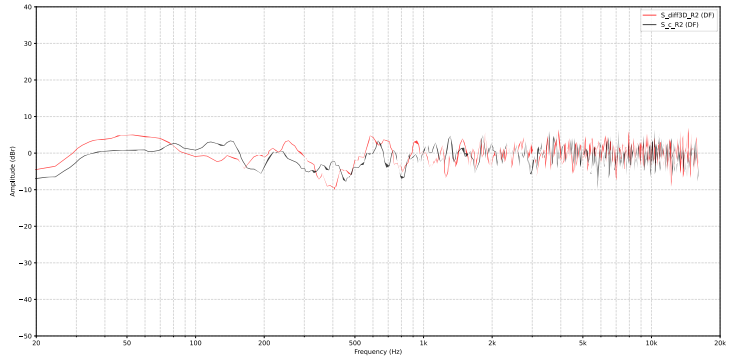


Figure 8.5 Frequency response plot comparison between S_{diff3D} and S_c from receiver 2

All the other graphs can be found from appendix H to appendix N.

8.5 Ambisonic analysis

Analyses based exclusively on omnidirectional measurements and classic acoustic parameters are inadequate for a comprehensive description of the acoustic behaviour of a room. In order to gain more detailed spatial information about the perceived sound field, it is therefore useful to also analyse the directional distribution of energy decay over time.[27,52]

The Odeon software facilitates the generation of third-order ambisonic impulse responses of the simulated models. These responses, when processed using the same procedure employed for the IRs measured in the real environment using the O-Zylia microphone array, enable a comparison of the directional

characteristics of the simulated and measured sound field. This offers a more comprehensive perspective on the acoustic behaviour of the environment.

8.5.1 Rose decays

The tables present a comparison between the lateral, median, and transverse sections of the rose decay, calculated from the ambisonic impulse responses measured in the real environment and those obtained from simulations in Odeon for different configurations of the room's acoustic model. The rows of these tables illustrate the outcomes of the measured model and the various configurations of the simulated models:

- M, which is the measured model.
- S_c, complex model, in which the diffusers are modeled in 3D with their complete geometry.
- S_s, simplified model, in which the diffusers are represented as two-dimensional elements.
- S_{bass}, model of the empty room with only bass traps.
- S_{empty} , model of a completely empty room, without acoustic treatments.
- $S_{diff_{3D}}$, empty room with only 3D-modeled speakers
- $S_{diff_{2D}}$, empty room with only the speakers represented as two-dimensional elements.
- S_{refl}, empty room with only reflective panels on the ceiling

The columns represent the different reception positions considered in the environment (Receiver 1, Receiver 2, Receiver 2 right, Receiver 2 left, Receiver 3, and Receiver 4).

The comparison with M demonstrates that models lacking acoustic treatments, such as S_{empty}, yield a more uniform and isotropic energy distribution, suggesting a less directionally structured reverberant field. The introduction of specific acoustic elements has been demonstrated to progressively modify this behaviour.

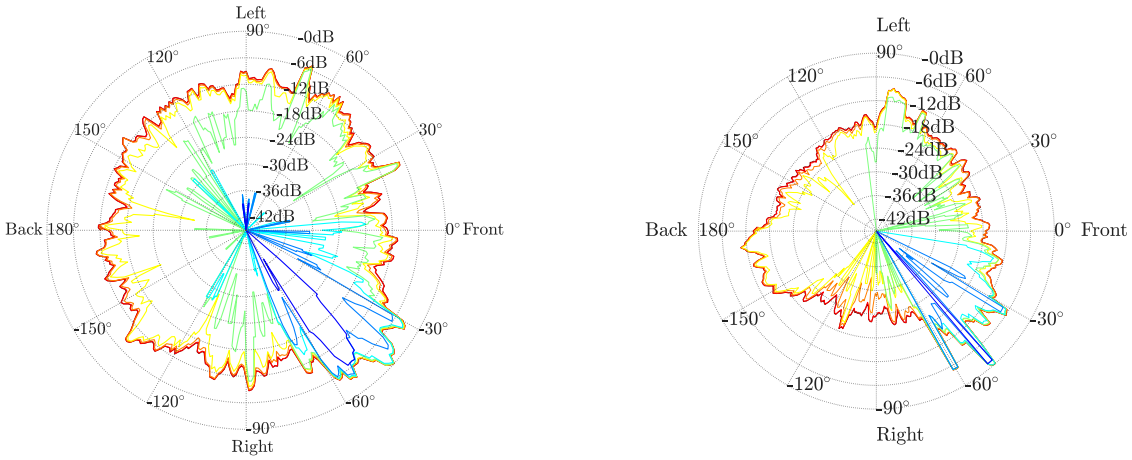


Figure 8.6 lateral section of the receiver R2 of the measured model M and the simulated empty room

The presence of diffusers has been shown to have a significant impact on the spatial structure of decay. In the $S_{diff_{3D}}$ and $S_{diff_{2D}}$ models, a redistribution of energy in different directions can be observed, which tends to bring the simulated behaviour closer to the measured one. Therefore, as demonstrated in the analysis of acoustic parameters, the spatial analysis also demonstrates that diffusers are the most effective element for the acoustic treatment of the environment. However, a comparison between the two configurations suggests that three-dimensional modelling of the speaker geometry allows for a more realistic reproduction of the angular dispersion of sound energy than their two-dimensional representation.This is particularly evident when observing the side section of microphone positions three and four, which are closest to the QRD panel and cylindrical panels, both in the $S_{diff_{2D}}$ and $S_{diff_{3D}}$ configurations and in the complete ones. As evidenced by the more diffuse and homogeneous redistribution of the very first and final reflections, S_c and S_s, as $S_{diff_{2D}}$ and $S_{diff_{3D}}$, demonstrate a distinct pattern. In the other positions of $S_{diff_{3D}}$, the rose decay also tends to have homogeneous contours and fewer concentrated directional peaks, which indicates the angular redistribution of the reflected energy. This verifies the standard functionality of a QRD diffuser, which serves to disrupt the specularity of reflections.[56] Conversely, within the $S_{diff_{2D}}$ model, the behaviour is more akin to that of an irregular reflective surface, as evidenced by the reduced homogeneity exhibited by the rose decay.

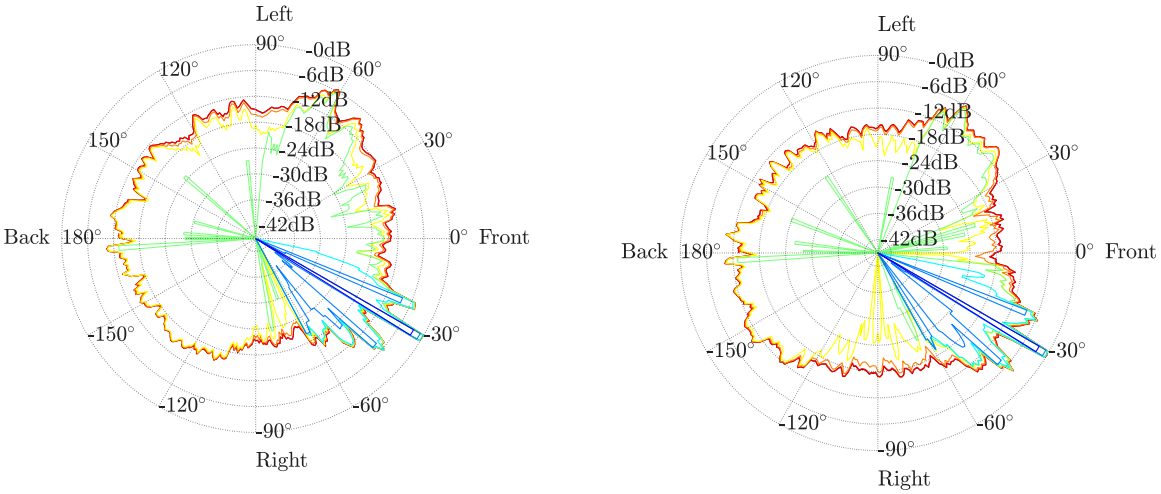


Figure 8.7 lateral sections of the rose decays of the receiver R3 of S_{diff3D} and S_{diff2D} on the right

The complex model S_c, incorporating the complete acoustic treatment system with three-dimensionally modelled diffusers, demonstrates the most significant overall similarity to the measured sound field. This suggests that a more detailed geometric representation is essential for a more precise description of the directional distribution of energy decay.

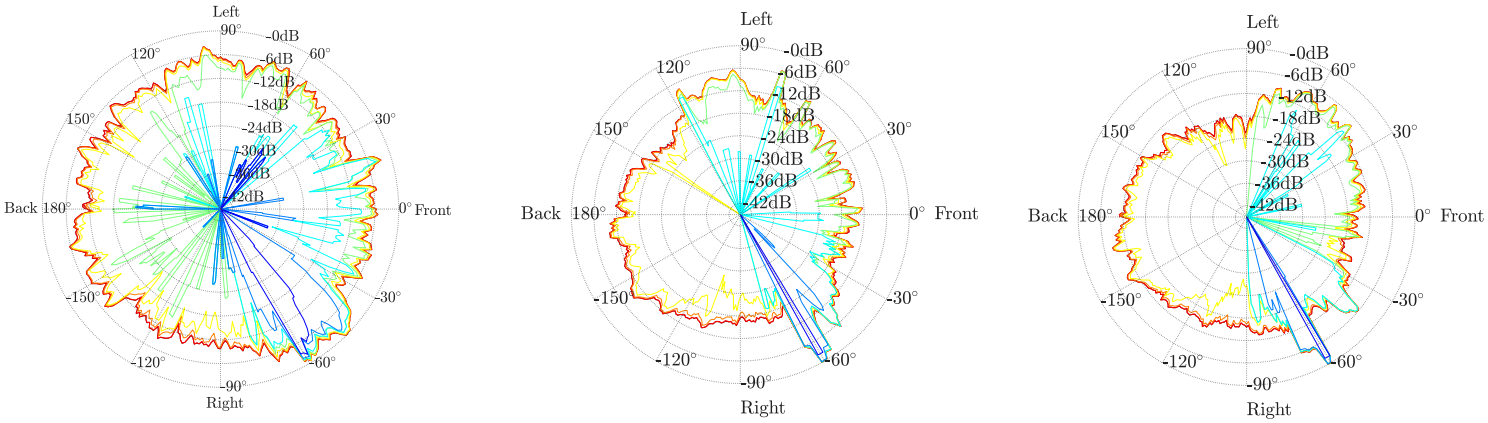


Figure 8.8 lateral section of the measured model, the complex S_c and the simplified one S_s

Notwithstanding this, the simulated rose decays are, in general, more regular and less irregular than the measured ones. This behaviour is characteristic of simulations based on geometric acoustics, in which the scattering models tend to produce more uniform energy distributions than the complexity of the real sound field. Moreover, analysis of the primary reflection directions facilitates the identification of these contributions as originating from the physical elements present in the room. For instance, by observing the cross sections of the position of receivers 2, 2 left, and 2 right, a dominant lateral reflection is apparent. This lateral reflection is more pronounced in some models than in others.

Another manifestation of the room's composition is evident in Figure 8.9 the spatial acoustic reflections, as observed in the first microphone position, which exhibits a more pronounced component of early reflections in an angular portion. This suggests the presence of an obstruction, which, in this case, is the wooden bench situated beneath the glass that overlooks the studio control room.

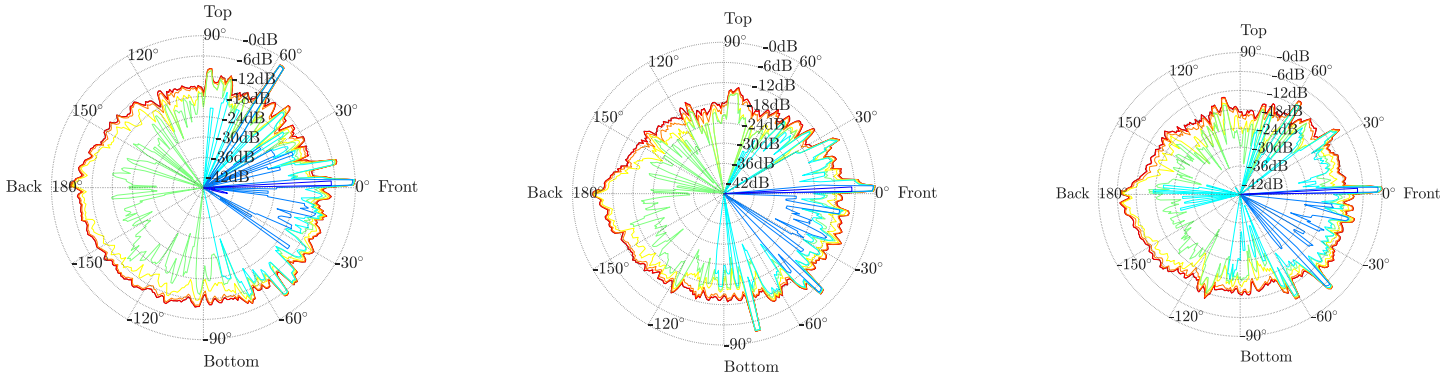


Figure 8.9 ambisconic IR median section from receiver R1 for the complex model, the empty room and the one just with the basstraps

8.6 Stage acoustic simulation with S_s and S_c

8.6.1 Set up on Odeon

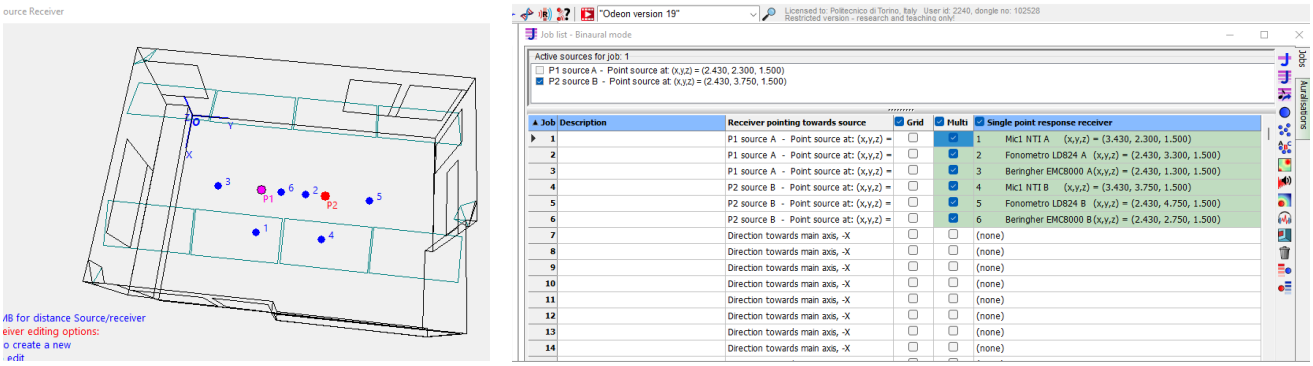


Figure 8.10 Setup of the acoustic simulation in Odeon: geometrical model of the room with source and receiver positions (left) and job configuration interface showing the active sources and receivers used for impulse response calculations (right).

The positions of the sources and receivers were then rearranged in order to reproduce exactly the two measurement configurations previously described. This adjustment was deemed necessary to ensure a consistent comparison between the measured acoustic data and the simulated conditions. The ray-tracing calculation was therefore performed by activating Source A and pairing it with the receivers belonging to Setup A. The same procedure was then repeated for Source B, associated with the receivers of Setup B, while ensuring that the geometric model of the room remained unchanged. The configuration illustrated in Figure 8.10 was subsequently replicated for the complex model, in which the diffusers were explicitly modelled in three dimensions. This allowed a more detailed representation of their scattering behaviour while maintaining identical source–receiver geometries across the different simulation scenarios. In the subsequent data processing stage, the binaural impulse responses (IRs) generated by Odeon from the aforementioned simulations were exported and analysed using the Aurora plug-in for Audacity. This apparatus facilitates the extraction of acoustic parameters from the impulse responses. In particular, the stage acoustics parameters were computed and used for the comparative analysis between the different simulated configurations and the measured acoustic condition.

8.6.2 Comparison with the Objective measurements

The following comparison yields a number of noteworthy results. Firstly, it is evident that the behaviour of the objective measurements is maintained by the simulation, irrespective of whether the model is complex or simplified. The geometric model is acoustically consistent with reality, successfully reproducing the relationship between early and late energy, which shows a very pronounced minimum at 250 Hz, moderately negative values between 500 Hz and 4 kHz, and stabilization at high frequencies. In contrast to the observations made in the simulation of the other measurement type, it is evident that the complex model does not systematically exhibit a closer proximity to the measured data in comparison to the simplified model. As the frequency descends below the range of 125 Hz to 250 Hz, a marked increase

in the discrepancy between the simulation and the measurement is observed. It is important to note that this latter characteristic is normal when one considers that, at low frequencies, room modes, wave phenomena and interference dominate. These phenomena are not well represented by geometric- acoustic models.[26]

The results reveal a characteristic transition between two acoustic regimes. In the low-frequency range the behaviour is dominated by modal effects, resulting in larger deviations between measured and simulated parameters and lower values of the ST_{late} and the ST_2. Above approximately 500 Hz the parameters become significantly more stable, indicating a transition towards a more diffuse sound field. This behaviour is typical of small treated rooms and confirms that the acoustic design provides controlled low-frequency behaviour while maintaining useful early reflections for performer support.

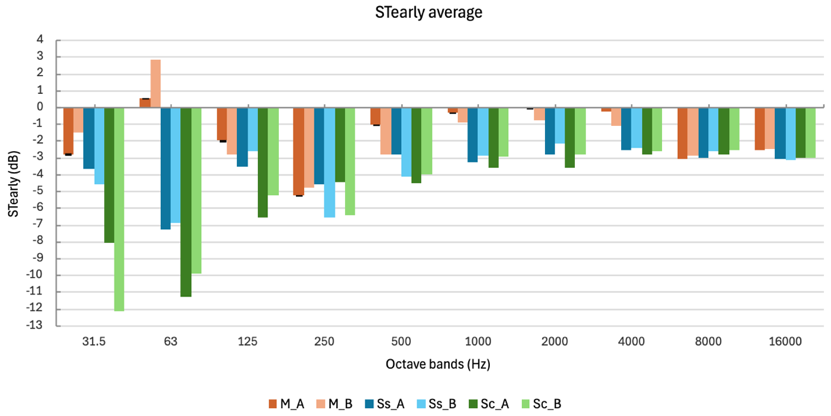


Figure 8.11 comparison with the ST_{early} averaged over the three receiver of the source setup A and B between the measured one with the complex model simulation and the simplified model simulation

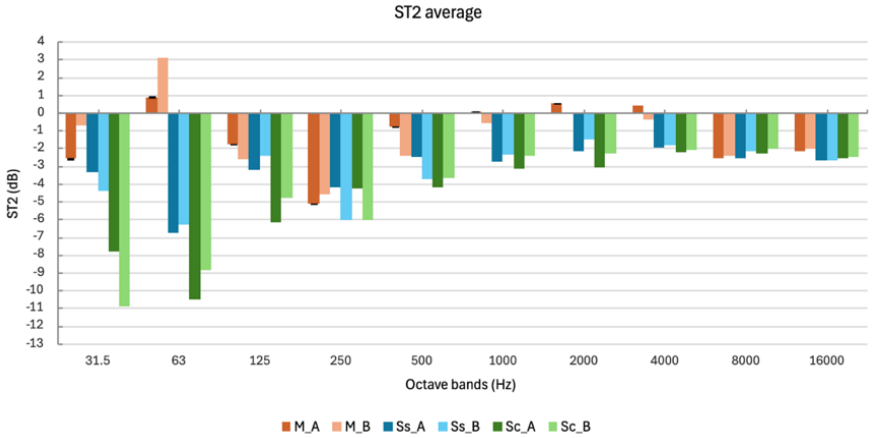


Figure 8.12 comparison with the ST_2 averaged over the three receiver of the source setup A and B between the measured one with the complex model simulation and the simplified model simulation

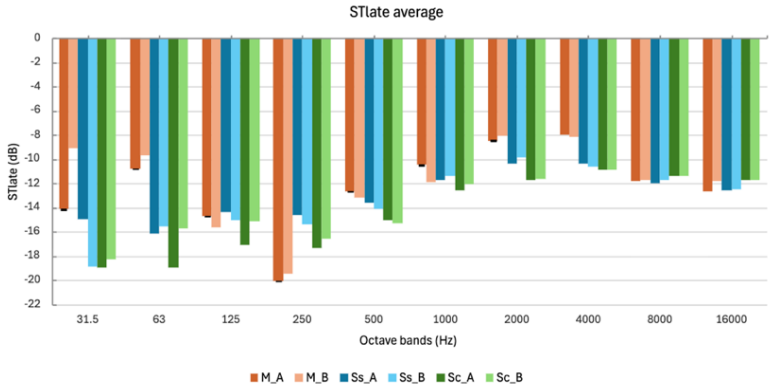


Figure 8.13 Figure 8.18 comparison with the ST_{late} averaged over the three receiver of the source setup A and B between the measured one with the complex model simulation and the simplified model simulation

Since the stage acoustic is an objective measurements of how musicians can hear their self playing or can hear each other in an ensamble situation[30], the key metrics are quietly influenced by their position in the acoustic space. In the case of our room the setup A has more flat early support, while setup B has a more late support response from the room. This peculiarity is more evident looking at the the receiver-wise breakdown, which can be found in Appendix Q.

9 Conclusions

This study analyzes the acoustic behavior of the recording studio housed at SERMIG in Turin through an integrated approach based on in situ measurements, acoustic parameter analysis, and numerical simulation. In the first phase, a series of measurements was carried out to determine the sound power level G of the room, assess speech intelligibility, and extract the main objective acoustic parameters from the measured impulse responses.
These include EDT, T_{20}, T_{30}, D_{50}, C_{50} and C_{80}, which describe the energy decay and temporal distribution of sound energy in the environment. In addition, stage acoustic measurements were taken to obtain the ST_{early} and ST_{late} values.

To further investigate the directional behavior of the sound field, a Spatial Ambisonic Impulse Response was also acquired using the O-Zylia array, allowing for an immersive analysis of the propagation and decay of energy in space.
Based on the measurements taken, a virtual reconstruction of the environment was subsequently developed and implemented in both the Odeon acoustic simulation software and the Blender environment using the VI-Suite AuVi add-on. Two levels of modeling were created: a simplified model that represents the scattering of the diffusers in a parametric way and a more detailed model that reproduces the geometric scattering through explicit three-dimensional modeling of the diffusers.
The simulations obtained with the two platforms were compared with each other and with the measured data, with the aim of evaluating the ability of the different models to reproduce the actual acoustic behavior of the room. Finally, different configurations of the model in Odeon were explored to analyze the acoustic contribution of the individual elements present in the space and understand their role in the distribution of sound energy.

Currently, there are no established guidelines that precisely define the optimal values for acoustic parameters in a live recording room. In this context, the SERMIG studio represents an interesting case study, not only for the objective measurements of acoustic parameters, but also for the stage acoustics analyses carried out in the space. These analyses, normally applied in theaters and concert halls, are oriented towards the sound perception of musicians on stage. Their application in a recording studio suggests a possible design approach to acoustic correction that is more attentive to the musician's perception in the space, and not exclusively to the quality of the signal perceived by the sound engineer in the control room. A significant result that emerged from the simulations conducted with Odeon concerns the greater reliability that can be achieved through a simulation with geometric scattering compared to a representation oriented towards parametric scattering. In VR simulations, it is the three-dimensionality and accuracy of the geometry that best represents the spatial distribution of sound energy within the real environment.

At the same time, the analysis of the AuVi features integrated into the Vi-Suite add-on for Blender developed by Ryan Southall was useful for understanding its limitations and valuable potential.

This tool, based on open-source software for three-dimensional modeling and computer graphics, represents an interesting attempt to extend Blender's capabilities to the field of energy and acoustic simulation.
The results obtained in this thesis highlighted some limitations when compared to established software developed for architectural acoustic analysis, such as Odeon. However, the development prospects for future versions of the add-on suggest that tools such as AuVi may in the future represent solutions capable of integrating the advanced three-dimensional modeling capabilities offered by Blender with increasingly mature tools for acoustic simulation. Simulating the acoustic contribution of individual treatment elements within the virtual model allows for predictive evaluation of the effectiveness of different correction strategies. This approach represents a design tool of great interest to acoustic engineers and designers, allowing them to test and optimize interventions before implementing them in the real environment. Finally, most of the measurements and simulations conducted during this thesis were accompanied by an immersive and spatial analysis of the sound field. This type of approach allows objective acoustic parameters to be integrated with a perceptual representation of the environment, constituting one of the most advanced tools currently available for simulating human perception of sound within a room in virtual reality environments.

9.1 Future works

A promising avenue for future research concerns a more in-depth investigation of stage acoustic parameters within the context of recording studios. Although these parameters have been extensively researched in concert halls and theatres, their systematic application to recording environments remains relatively unexplored. Stage acoustic metrics such as ST_1(or ST_{early}), ST_{late} and ST_2 , originally introduced to quantify the acoustic support perceived by musicians on stage, are typically employed in the analysis of orchestral platforms and performance spaces.[20,26]

Nevertheless, analogous considerations could prove to be of great significance in the context of recording studios, wherein musicians perform in the absence of click tracks or headphones, relying instead on the inherent acoustic feedback of the environment. A meaningful extension of this research would therefore involve identifying recording environments designed for ensemble interaction, particularly those used for acoustic genres in which performers rely on the room response to maintain musical cohesion. Following the selection of this subset of studios, it would be possible to analyse spaces already perceived by musicians as acoustically successful and use them as reference case studies. A preliminary investigation into the stage acoustic parameters was conducted, with the objective of establishing reference ranges for the ST_{early} , ST_{late} and ST_2 . These ranges were then utilised as guidelines for the design or evaluation of future recording rooms.

Another salient development pertains to the evolution of acoustic simulation tools. The AuVi add-on, included in the VI-Suite for Blender (developed by Ryan Southall), represents a particularly innovative approach to acoustic simulation within an open-source environment. In a cultural and technological

landscape increasingly characterised by home recording and user-generated content, the availability of accessible tools capable of providing meaningful acoustic predictions could significantly broaden the use of acoustic simulation beyond specialised professional contexts. For this reason, it would be valuable to reiterate the comparison presented in this work using more mature versions of the software and to benchmark its performance against well-established simulation platforms such as Odeon, EASE and Treble. These have already been extensively validated in the literature. [39,52] From a methodological perspective, future studies could also benefit from integrating classical acoustic parameters with spatial analysis techniques, such as ambisonic impulse responses and immersive virtual reality simulations. Recent research suggests that spatial impulse responses and binaural or ambisonic representations can provide a more detailed description of the spatial distribution of acoustic energy and of the directional characteristics of reflections. [36,57] It is therefore hypothesised that these approaches may support a more perception-oriented interpretation of room acoustics, complementing the information obtained from traditional monaural impulse response parameters.

A crucial direction for future work involves conducting systematic perceptual studies with musicians. The utilisation of structured listening tests and performance-based evaluations would facilitate the collection of qualitative and quantitative feedback on the influence of different acoustic environments on musical interaction, timing, and ensemble cohesion. The implementation of such investigative methodologies has the potential to facilitate a more seamless integration between objective acoustic measurements and the subjective experience of performers. This, in turn, could contribute to the design of recording studios that emulate the acoustic conditions frequently associated with the most successful concert halls and performance spaces.[26] In this perspective, the design of recording studios could progressively evolve from a purely control-room-oriented approach toward a performer-oriented acoustic design, where the acoustic environment is shaped not only to optimize recording quality but also to enhance the interaction, communication, and expressive freedom of musicians during performance. By combining objective acoustic measurements, spatial analysis techniques and perceptual evaluations, future research may contribute to defining new design strategies capable of creating recording environments that support both technical accuracy and musical inspiration.

Bibliography

1. *Acoustics, Measurement of Room Acoustic Parameters. Part 1 : : Performance Spaces (ISO 3382-1:2009)*. BSI Standards, 2009; 2012.
2. *AMERICAN NATIONAL STANDARD Criteria for Evaluating Room Noise*.
3. Kuttruff H. *Room Acoustics, Fifth Edition*.
4. Hansen CH. *Understanding Active Noise Cancellation*. Taylor & Francis eLibrary; 2003.
5. Von Lauer-Münchhofen J, Von Berg M, Steffens J. *Applicability of Common Room Acoustic Parameters for Music Recording Spaces in Recording Studios*.
6. Tavelidou E, Foteinou A, Spyridis C. *Acoustic Study of the Live Room*.
7. Pagine 1029 - 1048 (studi di registrazione) ACUSTICA_FONDAMENTI ED APP - spagnolo_UTET.
8. *Acoustics, Measurement of Room Acoustic Parameters. Part 1 : : Performance Spaces (ISO 3382-1:2009)*. BSI Standards, 2009; 2012.
9. *BSI Standards Publication Acoustic Quality Criteria for Music Rehearsal Rooms and Spaces*. 2025.
10. Şaher K, Özgencil Y, Khoshkholghi S. Acoustic assessment of four music rehearsal rooms in accordance with ISO23591 standard. *Applied Acoustics*. 2025;240. doi:10.1016/j.apacoust.2025.110878
11. Eaton J, Gaubitch ND, Moore AH, Naylor PA. Estimation of Room Acoustic Parameters: The ACE Challenge. *IEEE/ACM Trans Audio Speech Lang Process*. 2016;24(10):1681-1693. doi:10.1109/TASLP.2016.2577502
12. Von Lauer-Münchhofen J, Von Berg M, Steffens J. *Applicability of Common Room Acoustic Parameters for Music Recording Spaces in Recording Studios*.
13. *Acoustics, Measurement of Room Acoustic Parameters. Part 2 : Reverberation Time in Ordinary Rooms*. BSI Standards; 2008.
14. Del Solar Dorrego F. *A STUDY OF THE JUST NOTICEABLE DIFFERENCE OF EARLY DECAY TIME (EDT)*. Vol 40. 2018.
15. Rindel JH. *Room Acoustic Measurement Methods in the Past, Present and Future, Including the Importance of the ISO 3382 Series*. https://www.researchgate.net/publication/381462119
16. Newell P. General Requirements and Common Errors. In: *Recording Studio Design*. Elsevier; 2012:1-12. doi:10.1016/b978-0-240-52240-1.00001-6
17. Farina A, Commins DE, Prodi N. *Experimental Analysis of the Acoustical Behaviour of Musikverein in Concert and Ballet Configurations*.
18. Von Lauer-Münchhofen J, Von Berg M, Steffens J. *Applicability of Common Room Acoustic Parameters for Music Recording Spaces in Recording Studios*.
19. Tavelidou E, Foteinou A, Spyridis C. *Acoustic Study of the Live Room*.
20. Michael Barron. *Auditorium Acoustics and Architectural Design*.

21. Beranek LL. *Concert Halls and Opera Houses : Music, Acoustics, and Architecture*. Springer; 2004.
22. *Acoustics, Measurement of Room Acoustic Parameters. Part 1 : : Performance Spaces (ISO 3382-1:2009)*. BSI Standards, 2009; 2012.
23. Kalkandjiev Z;, Weinzierl S. The Influence of Room Acoustics on Solo Music Performance: An Empirical Case Study. *Acta Acustica united with Acustica*. 2013;99(3):433-441. doi:10.14279/depositonce-15277
24. Von Békésy, Georg. *Feedback Phenomena between the Stringed Instrument and the Musician*. (1968).
25. Kato K, Ueno K, Kawai K. Musicians' adjustment of performance to room acoustics, part III: Understanding the variations in musical expressions. *Journal of the Acoustical Society of America*. 2008;123.
26. A. C. Gade. Musicians acoustic conditions Part I Methods and laboratory experiments Acustica1989. *Acoustica*. Published online 1989.
27. Tervo S, Pätynen J, Lokki T. *Spatial Analysis of Concert Hall Impulse Responses*. 2013.
28. Pätynen J, Tervo S, Lokki T. Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses. *J Acoust Soc Am*. 2013;133(2):842-857. doi:10.1121/1.4770260
29. Shtrepi L, Astolfi A., D'Orazio D., Prato A. Objective and perceptual evaluation of stage acoustics in concert halls. *Applied Acoustics*. Published online 2017.
30. Shtrepi L., Astolfi A. Stage acoustic parameters and musicians' perception. . *Journal of the Acoustical Society of America (JASA)*. Published online 2019.
31. Oppenheim AV., Schafer RW. *Discrete-Time Signal Processing*. Pearson; 2014.
32. Farina A. *Advancements in Impulse Response Measurements by Sine Sweeps*. 2007. www.aes.org.
33. M. R. Schroeder. Schro65-reverb. Published online December 14, 1964.
34. F. A. & PKC. *Master Handbook of Acoustics (6th Ed.)*. (McGraw-Hill., ed.). 2015.
35. Toole FE. *Sound Reproduction : Loudspeakers and Rooms*.
36. Farina A. *Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique*.
37. Kuttruff H. *Room Acoustics, Fourth Edition*.
38. *Acoustics, Measurement of Room Acoustic Parameters. Part 1 : : Performance Spaces (ISO 3382-1:2009)*. BSI Standards, 2009; 2012.
39. Vorlander. *Auralization*. Springer Berlin Heidelberg; 2008. doi:10.1007/978-3-540-48830-9
40. Jean-Francois B. *General Rights Efficient Low-Frequency Room Acoustic Modelling*.
41. Siltanen S, Lokki T, Savioja L. *Rays or Waves? Understanding the Strengths and Weaknesses of Computational Room Acoustics Modeling Techniques*. 2010.
42. Postma BNJ, Katz BFG. Creation and calibration method of acoustical models for historic virtual reality auralizations. *Virtual Real*. 2015;19(3-4):161-180. doi:10.1007/s10055-015-0275-3
43. Harris CM. *Handbook of Acoustical Measurements and Noise Control*. McGraw-Hill; 1991.

44. Knauf-Danoline-Data_sheet_Plaza_UK1-1.
45. Fuchs HV. *Schallabsorber Und Schalldämpfer : Innovative Akustische Konzepte Und Bauteile Mit Praktischen Anwendungen in Konkreten Beispielen*. Springer-Verlag; 2010.
46. Ua WA. *ABSORPTION COEFFICIENTS*. www.akustik.ua
47. Hansen CH. *Acoustic Absorbers and Diffusers*. Taylor & Francis eLibrary; 2003.
48. Christensen CL, Rindel JH. *A New Scattering Method That Combines Roughness and Diffraction Effects*.
49. Hansen CH. *Understanding Active Noise Cancellation*. Taylor & Francis eLibrary; 2003.
50. Ryan Southall. VI-SuiteView Messages An integrated suite of environmental analysis tools within Blender. September 23, 2024. Accessed November 17, 2025. https://blogs.brighton.ac.uk/visuite/2024/09/23/auvi-and-acoustics/
51. Kleiner Mendel, Tichy Jiri. *Acoustics of Small Rooms*. CRC Press; 2014.
52. Lynge Christensen C, Koutsouris G, Gi J. *ODEON Room Acoustics Software User Manual*. www.odeon.dk
53. Ryan. *Simulations and Visualisations with the VI-Suite*.
54. Sluyts Y, Kritly L, Vandenberghe A, Glorieux C, Rychtarikova M. Case study: the influence of model size on local room acoustic parameters in Odeon. In: *Building Simulation Conference Proceedings*. International Building Performance Simulation Association; 2022:2805-2812. doi:10.26868/25222708.2021.30938
55. Nguyen H. *THE EFFECT OF FURNITURE ON ROOM ACOUSTIC PARAMETERS AND ITS DEPENDENCE ON THE SOUND ABSORBING PROPERTIES OF THE CEILING*. 2023. www.akustik.lth.se
56. Khrystoslavenko O, Grubliauskas R. Theoretical predictions of sound scattering coefficient and sound diffusion coefficient from quadratic residue diffusers. In: Vilnius Gediminas Technical University; 2020. doi:10.3846/aainz.2019.007
57. Rindel JH, Shiokawa H, Christensen CL, Gade AC. Comparisons between computer simulations of room acoustical parameters and those measured in concert halls. *J Acoust Soc Am*. 1999;105(2_Supplement):1173-1173. doi:10.1121/1.425555
58. Hansen CH. *Understanding Active Noise Cancellation*. Taylor & Francis eLibrary; 2003.

Statement on the Use of Artificial Intelligence

Artificial intelligence-based tools were used during the preparation of this thesis solely as a support for language editing, grammar correction, and improvement of textual clarity.

The conceptual development of the work, the mathematical modeling, the implementation of the algorithms, the analysis of the results, and the scientific interpretations presented in this thesis were entirely carried out by the author.The author takes full responsibility for the content of the manuscript.

Appendix

A. IR graphs for each receiver

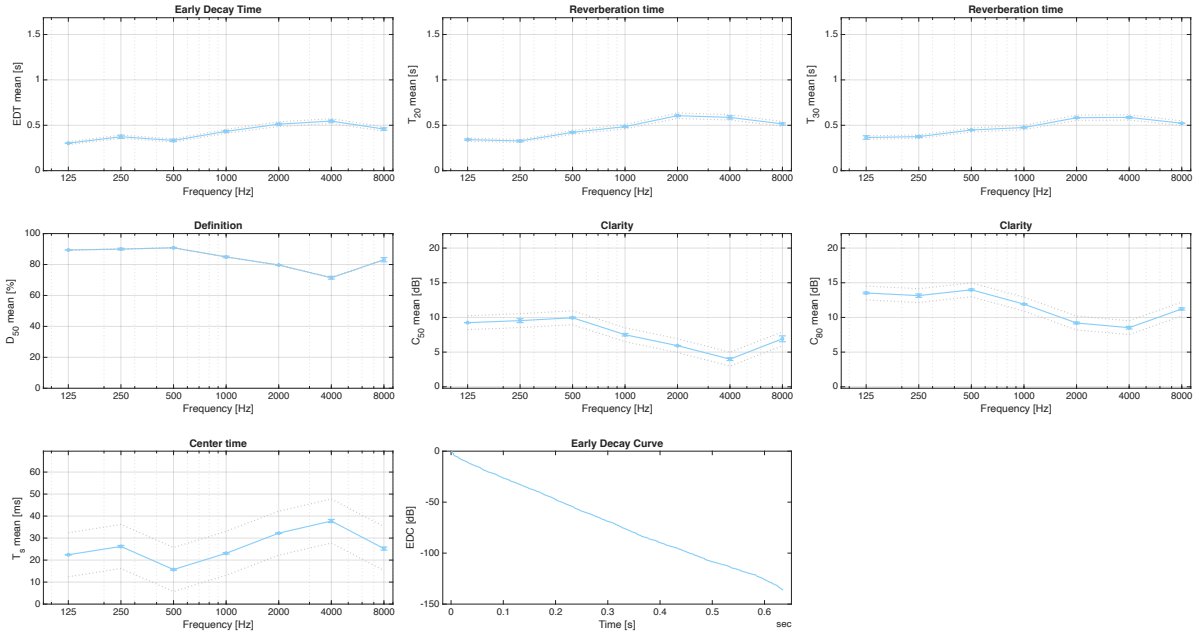


Figure 9.1 Room acoustic parameters over frequency for measurement position 1 (light-blue dots: receivers mean, grey dotted lines: JND) [ISO 3382]

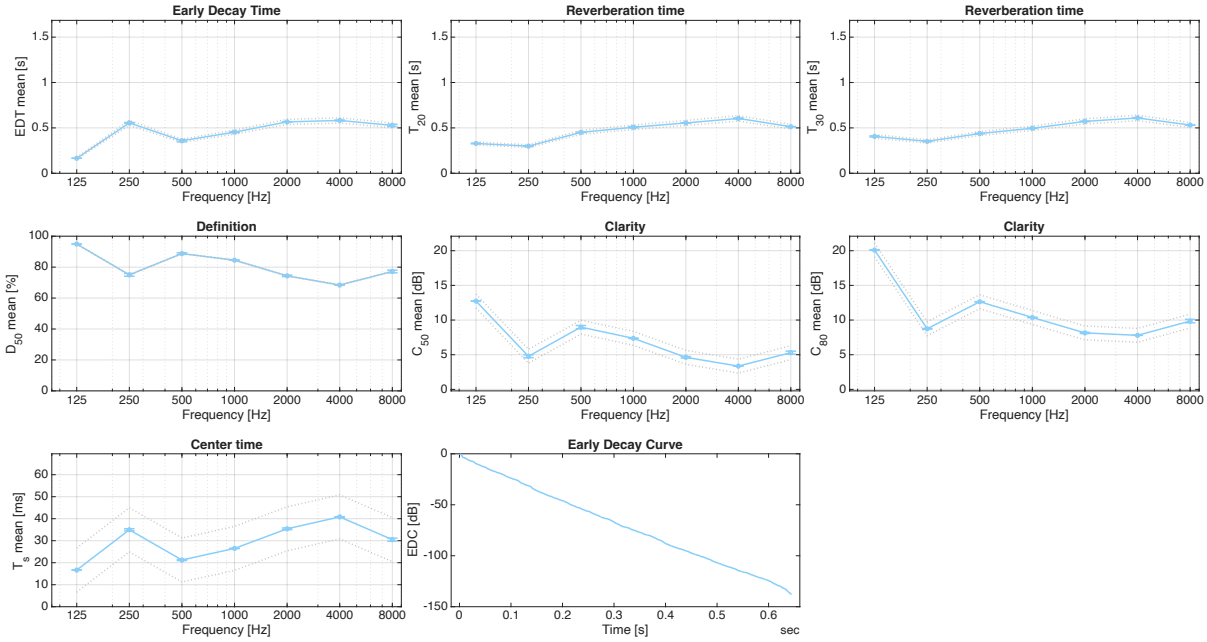


Figure 9.2 Room acoustic parameters over frequency for measurement point 2 (light-blue dots: receivers mean, grey dotted lines: JND) [ISO 3382]

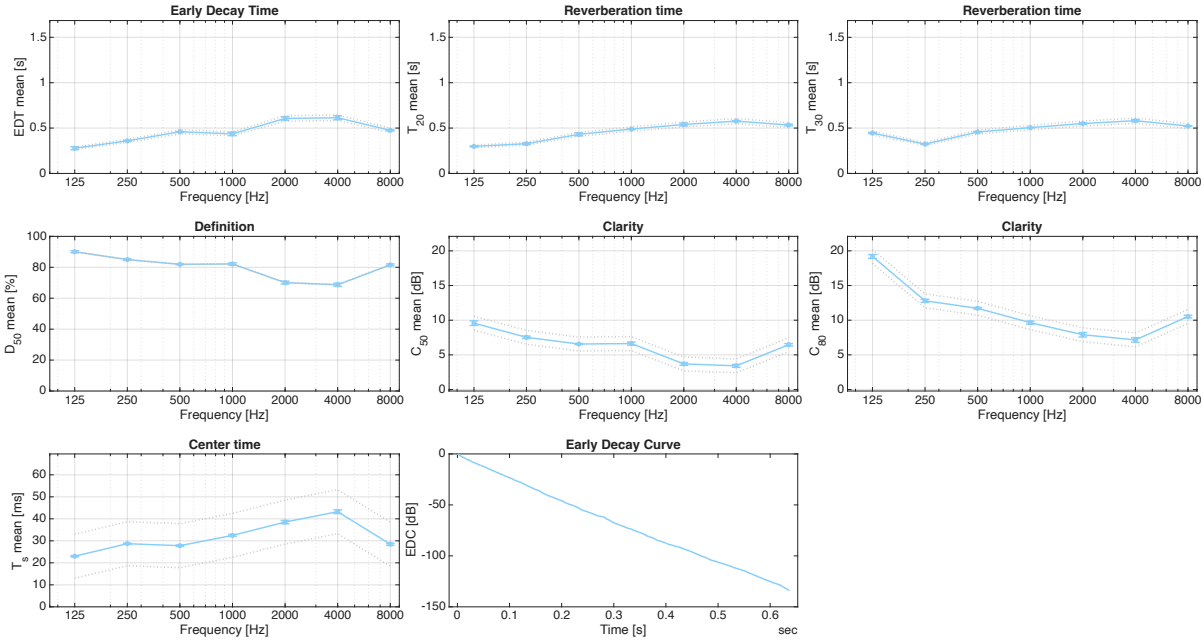


Figure 9.3 Room acoustic parameters over frequency for measurement point 2 left (light-blue dots: receivers mean, grey dotted lines: JND) [ISO 3382]

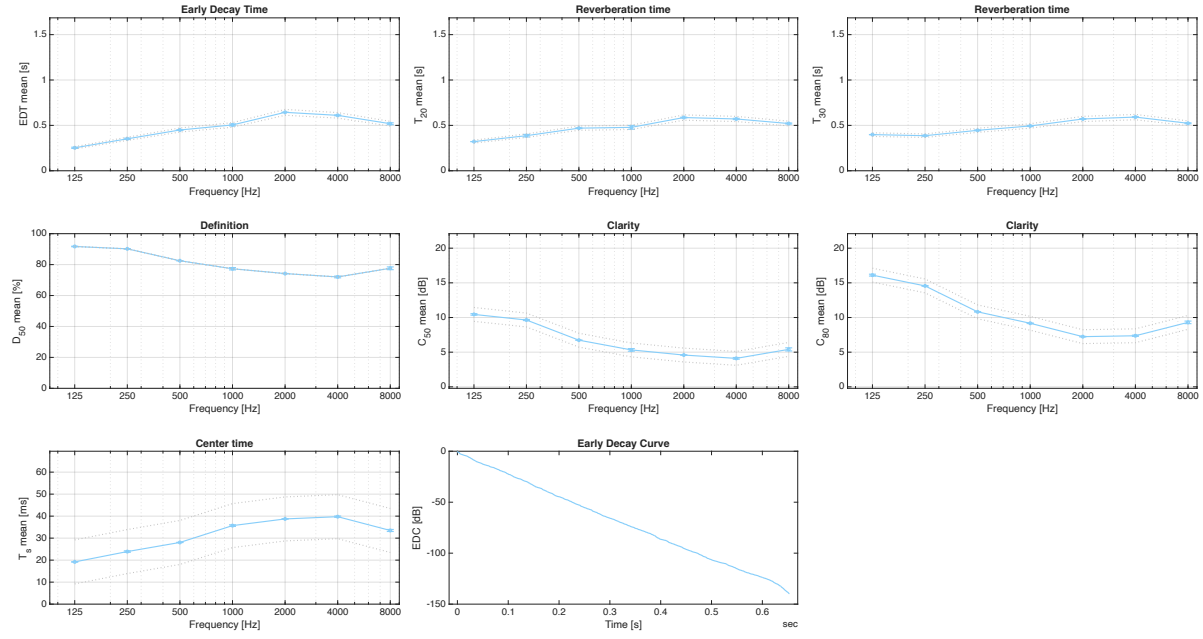


Figure 9.4 Room acoustic parameters over frequency for measurements point 2 right (light-blue line: receivers mean, grey dotted lines: JND) [ISO 3382]

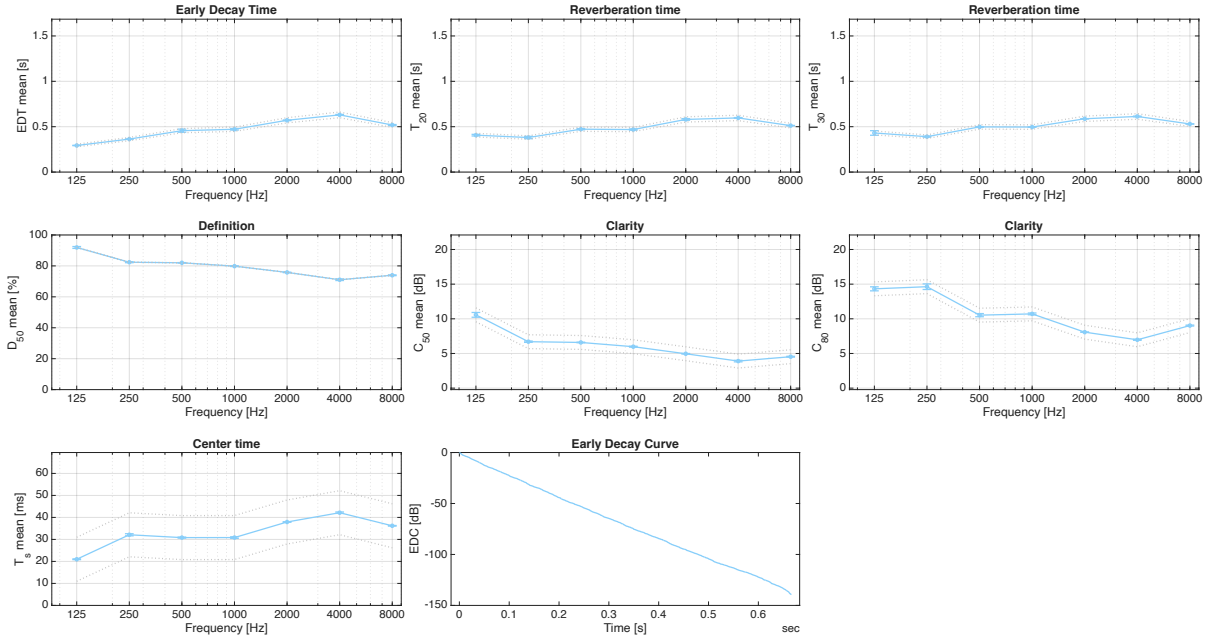


Figure 9.5 Room acoustic parameters over frequency for measurement point 3 (light-blue dots: receivers mean, grey dotted lines: JND) [ISO 3382]

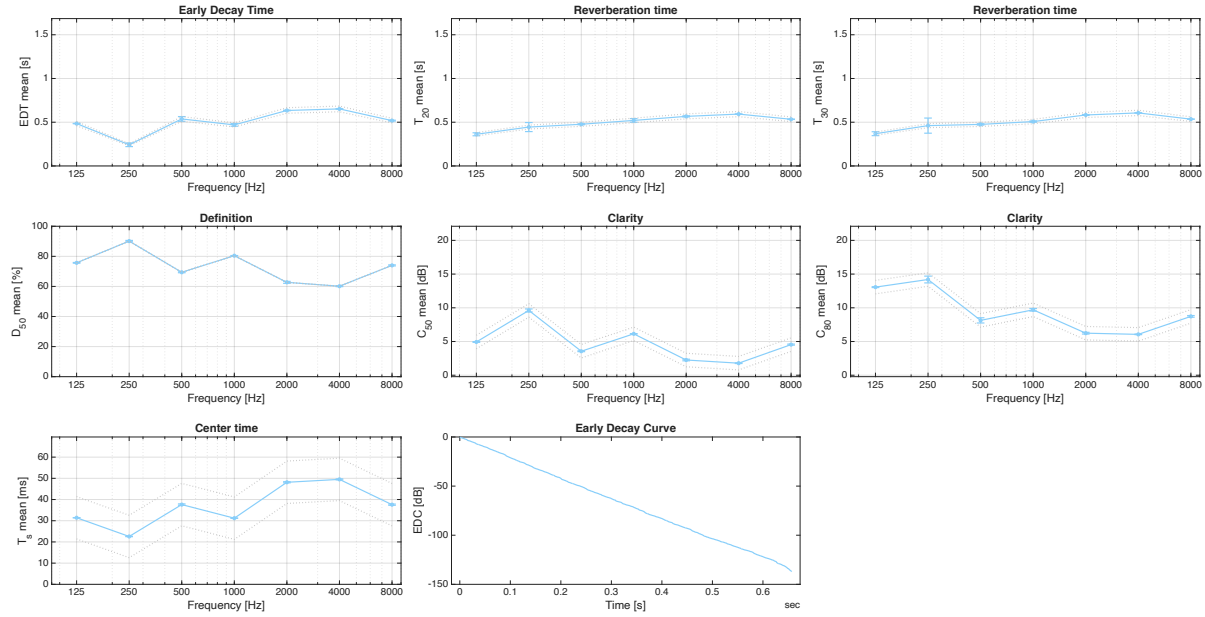


Figure 9.6 Room acoustic parameters over frequency for measurement point 4 (light-blue dots: receiver mean, grey dotted lines: JND) [ISO 3382]

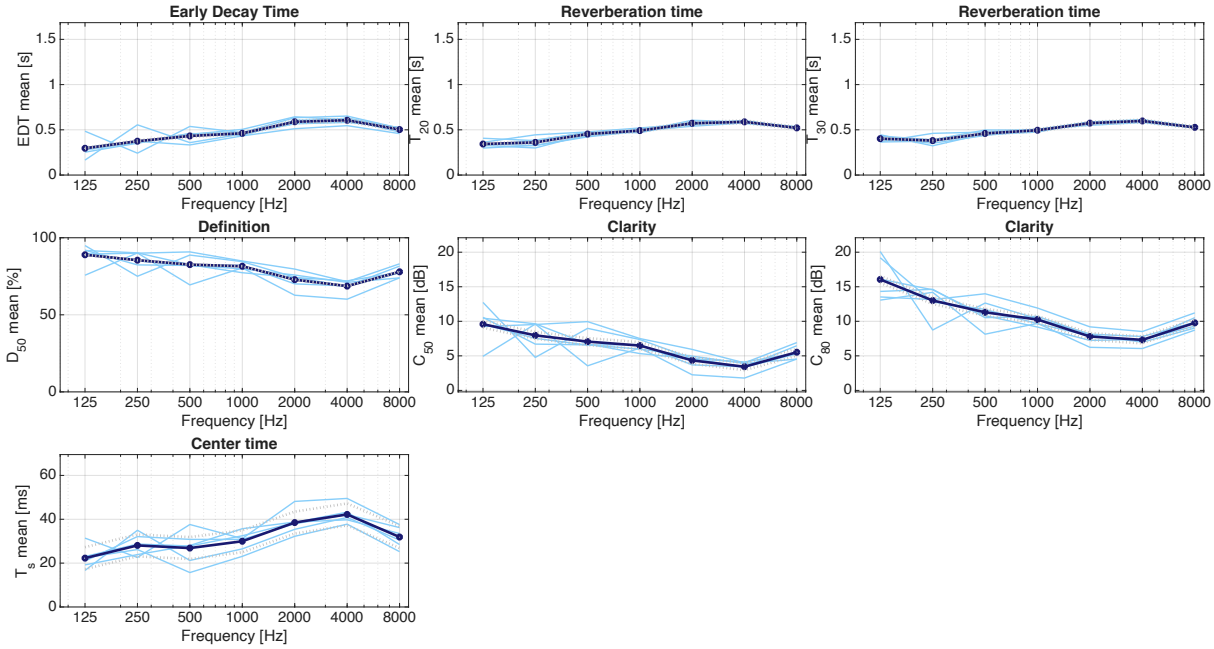


Figure 9.7 Overall mean values versus frequency (dark-blue line: receivers mean, light-blue lines: single receivers, grey dotted lines: JND) [ISO 3382]

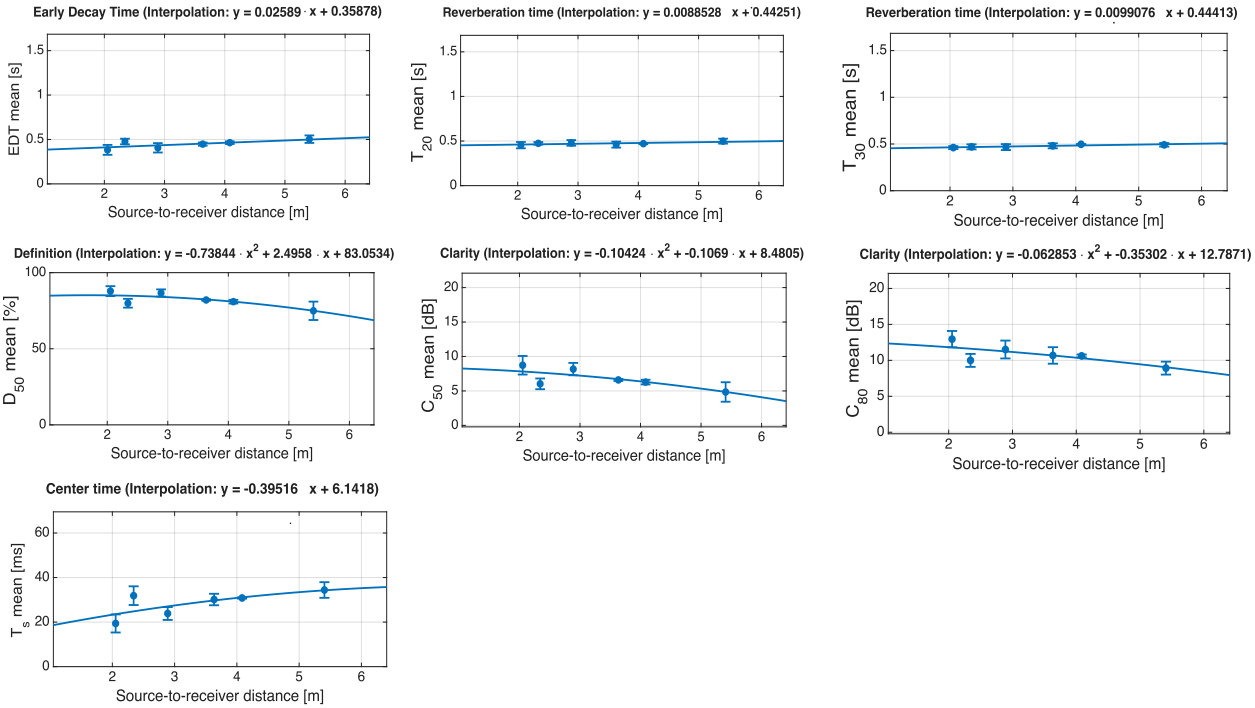


Figure 9.8 Overall meanvalues versus source-receiver distance (dark-blue dots: receivers mean, light-blue lines:interpolation) [ISO 3382]

B. Scattering coefficient in detail

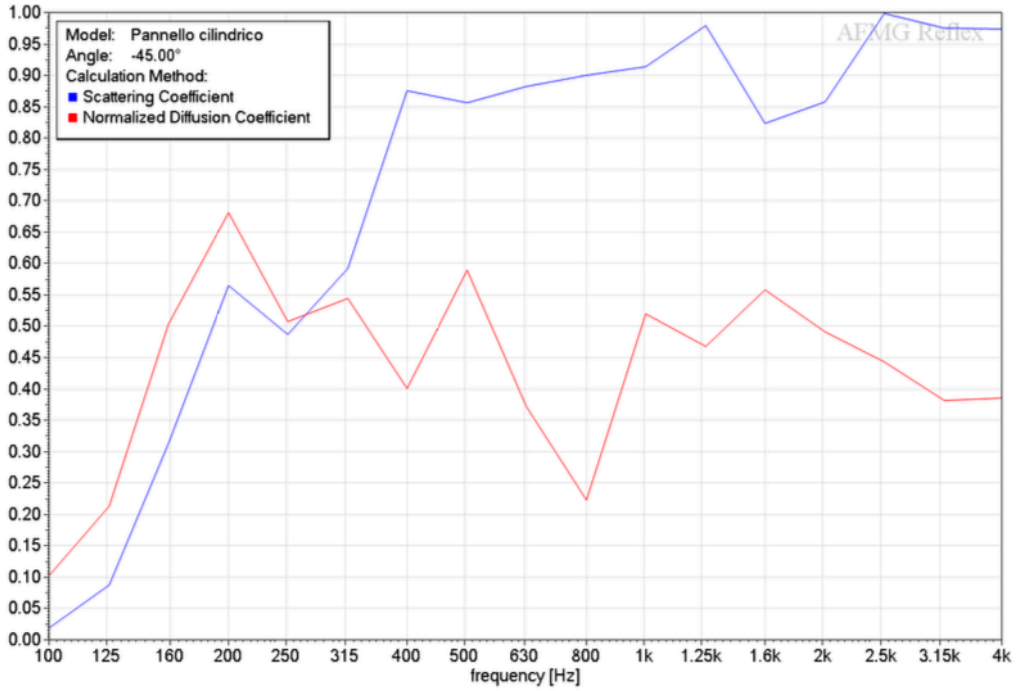


Figure 9.9 graph showing scattering and diffusion coefficients of the cilindrical panel according to ISO 17497-1 and ISO 17497-2 standards, calculated using a mathematical model

Table 9.1scattering coefficients averaged for the frequency bands of interest for the simulation

Coefficients	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
Scattering	0.075	0.5	0.86	0.92	0.85	0.97	0.97

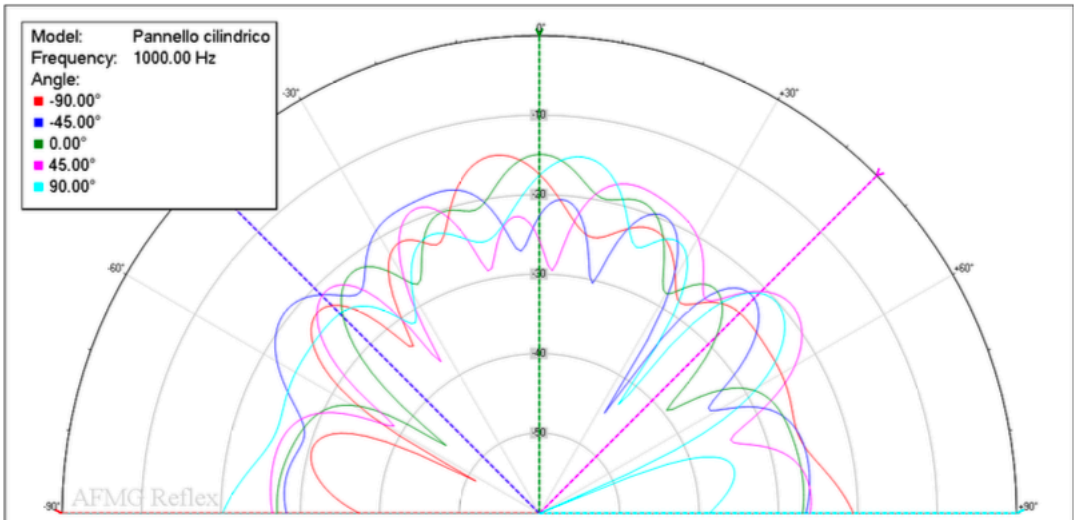


Figure 9.10 Spatial response of the panel for different angles at 1000 Hz, graph obtained through a mathematical model

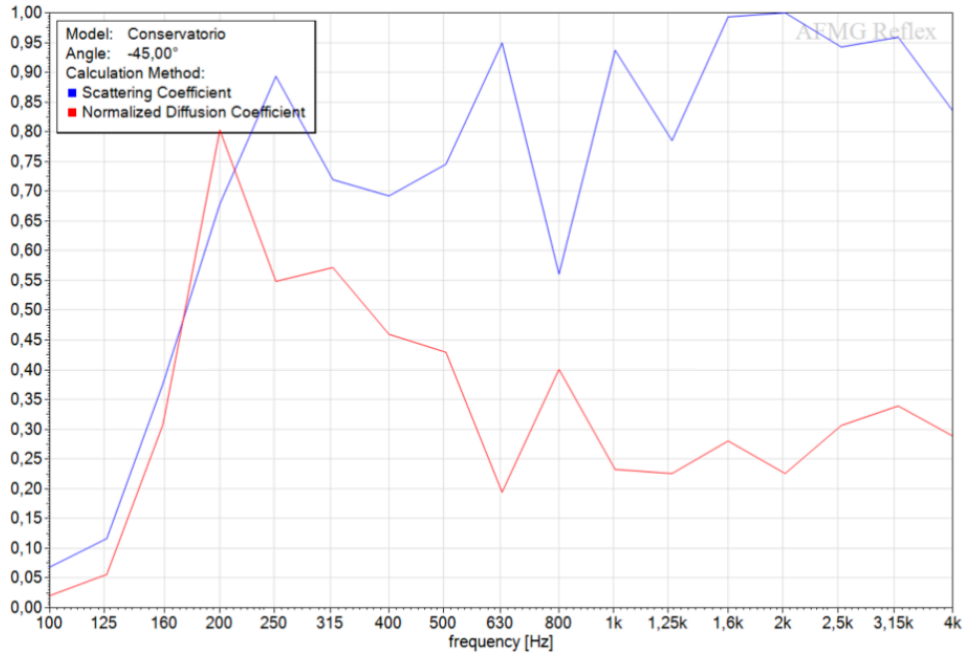


Figure 9.11 graph showing scattering and diffusion coefficients of the Schroeder diffuser according to ISO 17497-1 and ISO 17497-2 standards, calculated using a mathematical model

Table 9.2 scattering coefficients averaged for the frequency bands of interest for the simulation

Coefficients	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
Scattering	0.12	0.86	0.76	0.88	0.97	0.82	0.64

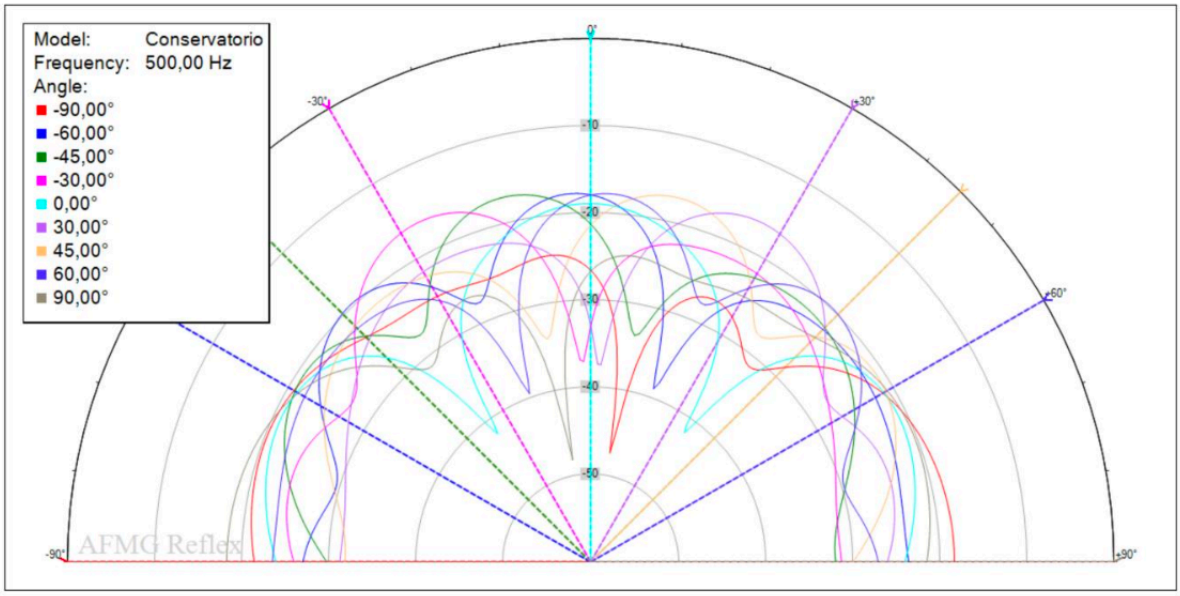


Figure 9.12 Spatial response of the panel for different angles at 500 Hz, graph obtained through a mathematical model.

C.Calibration steps for Odeon model

Table 9.3 Acoustic material properties adopted in the simulation model before calibration, including absorption coefficients across octave bands and scattering coefficients averaged over 500-1000 Hz, based on literature and manufacturer data.

Component	Material	Scattering coefficient (μ_s)	Absorption coefficient (α)						
			125 Hz	200 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
walls	Plaster. gypsum. or lime. smooth finish on lath (Harris, 1991)[43]	0.01	0.14	0.1	0.06	0.040	0.04	0.03	0.02
Stone walls	Brick. Unglazed (Harris, 1991)[43]	0.2	0.03	0.03	0.033	0.044	0.05	0.07	0.09
ceiling	Knauf Danoline Plaza. sound-absorbing forated plasterboard (manufacturer data)[44]	0.01	0.450	0.6	0.7	0.600	0.55	0.45	0.35
floor	Wood parquet in asphalt on concrete (Harris. 1991)[43]	0.01	0.042	0.042	0.074	0.063	0.063	0.074	0.084
windows	Double glazing. 2-3mm glass. >30mm gap (Fasold Winkler, 1976)[45]	0.01	0.15	0.05	0.03	0.04	0.02	0.02	0.02
studio window	Double glazing. 2-3mm glass. >30mm gap (Fasold Winkler, 1976)[45]	0.01	0.15	0.05	0.03	0.04	0.02	0.02	0.02
steel doors	Solid timber door (akustik.ua)[43,46]	0.01	0.140	0.100	0.06	0.08	0.1	0.1	0.1
bench	Wood parquet in asphalt on concrete (Harris. 1991)	0.01	0.04	0.04	0.07	0.06	0.06	0.07	0.08
cylindrical diffusing elements	Thin Plywood paneling (Ref. Dalenback. CATT)	0.89 *	0.42	0.21	0.1	0.08	0.06	0.06	0.06
Ceiling suspended diffusive panel	8mm wooden ceiling suspended panel (Cox & D'Antonio, 2003)[47]	0.01	0.05	0.08	0.08	0.1	0.1	0.1	0.05
Schroeder diffuser	Thin Plywood paneling (Ref. Dalenback. CATT)	0.87 *	0.05	0.08	0.08	0.1	0.1	0.1	0.05
Bass traps	pressed polystyrene EPS (Cox & D'Antonio, 2004)	0.01	0.01	0.01	0.02	0.02	0.035	0.06	0.07

The calibration process was initiated by entering the precise values of the materials and absorption and scattering coefficients into the designated sections of the software interface. In order to calibrate the model, an initial comparison was made between the T20 and the model, in order to ascertain the most appropriate way to make the T20 value curve tend towards that shown by the model as a function of frequency.
The initial values for the materials are delineated in the accompanying Table 9.3.

Subsequent to the initial simulation, the T20 curve of the simulation exhibited a significant discrepancy from the curve of the values measured on site for the same parameter, as illustrated in Figure 9.13.

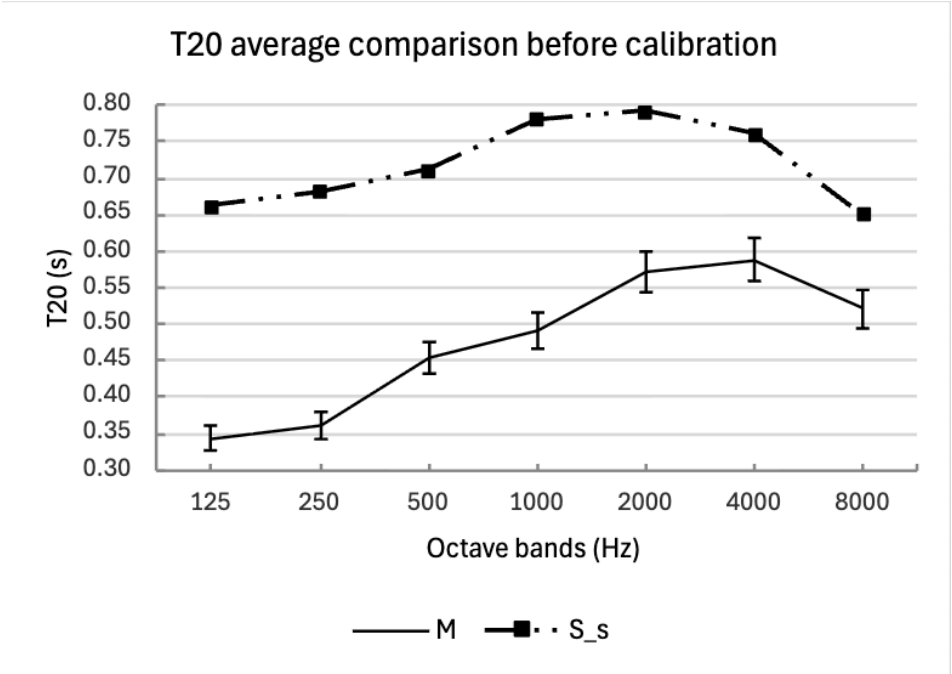


Figure 9.13 Comparison of octave-band T20 values between measurements (M) and simulations (S_s). Error bars indicate the variability of the measured data.

In order to reduce the reverberation time, it is necessary to increase the absorbency of the surfaces encountered by the sound inside the room. The most logical solution to this problem would be to increase the absorption coefficients of all materials.
These preliminary modifications were applied consistently to all materials across all octave bands, with the aim of preserving the original shape of their absorption curves.
In the initial two simulations, the absorption coefficients were augmented by 5% and 10%, respectively, with the objective of diminishing the reverberation time values. Nevertheless, the simulated room demonstrated higher levels of reverberation in comparison to the measured model, signifying that this adjustment proved to be inadequate.
To achieve more effective changes, it was decided to intervene on the largest surfaces, which are therefore the most influential from an acoustic point of view, namely the floor, walls, excluding the portion of unglazed stones, and ceiling. On these surfaces, the absorption coefficients were increased by 20% and

then by 25% across all frequencies compared to the initial values. Subsequently, the scattering coefficients of the surfaces exhibiting the greatest reflectivity, i.e. the walls, including the two stone walls, were increased from 0.05 to 0.3.

As the aforementioned interventions proved ineffective, a decision was taken to modify the absorption coefficients of the Schroeder diffuser materials and cylindrical panels, increasing them by 0.1 across all frequencies.

In consideration of the potential causes of the persistent discrepancy between the measured values of reverberation and the expected values, a hypothesis was formulated. This hypothesis proposed that, at the modelling level, certain elements characterised by high sound diffusivity had been overlooked. Examples of such elements include musical instruments present during the measurements. Of particular note was the presence of a grand piano, an organ, and an analog Leslie speaker. In order to take into account, at least parametrically, the diffusive effect introduced by these elements, the scattering coefficient of the floor was increased to 0.3.
In the latest attempts, the absorption coefficients of the cylindrical panels were increased by 0.1 across all frequencies. Concurrently, the absorption curve of the bass traps, windows, and glass was remodelled to engender a more analogous behaviour to that of a vibrating panel. In particular, an attempt was made to reproduce a resonant pattern that is characterised by an increase in absorption at low frequencies in the vicinity of the resonance frequency and a progressive reduction in effectiveness at high frequencies. This choice is motivated by the fact that these elements can behave as mass-spring systems, contributing to sound absorption mainly in the low-frequency region.[47]
With regard to the subsequent revisions, the two radical alterations that facilitated calibration with the measured model entailed the complete substitution of the floor material and the Schroeder diffuser material.
With regard to the flooring, given that the room has been constructed according to a box-in-a-box configuration, it was considered appropriate to hypothesise that the parquet was not directly laid on a concrete slab. It can be hypothesised that, given the nature of the construction method employed, the parquet flooring may be situated on a pre-existing intermediate layer. This layer, which can be traced back to the original construction phase, remains to be analysed in terms of its composition.
In order to take this hypothesis into account and make the surface more absorbent, the floor absorption coefficient curve was modified. The curve for parquet laid on asphalt was replaced with the one for parquet installed on counterfloor.
With regard to the Schroeder diffuser, the previously observed absorption curve, associated with thin plywood panelling, has been superseded by that of a one-dimensional quadratic residue diffuser with seven inlets. This new curve was derived from the text by Cox and D'Antonio and is referred to as "1D N = 7 QRD, design frequency = 500 Hz." [47] The term "quadratic residue diffuser" is employed to denote a particular acoustic diffuser, the function of which is to disperse sound in space, as opposed to reflecting it specularly. The 1D type of the model suggests that diffusion occurs in only one direction, either horizontally or vertically, due to the presence of a series of parallel wells with varying depths. The parameter N, which denotes the number of wells that constitute the fundamental module of the diffuser,

is determined by a specific mathematical law that governs the sequence of depths. Finally, the design frequency of 500 Hz indicates the centre frequency around which the diffuser is optimised, i.e. the frequency range for which the panel is most effective in diffusing sound energy. It is evident that the most recent modifications have resulted in the model demonstrating a greater proximity to the calibration target. Nevertheless, a divergence remains apparent in the low and mid-low frequency range, notably below 1000 Hz.

In order to further reduce this discrepancy, targeted action was taken on this frequency range, increasing the absorption coefficients of the floor, walls, and ceiling exclusively in that band. This choice enabled the enhancement of the model's response at the most critical frequencies without compromising its performance at medium-to-high frequencies.

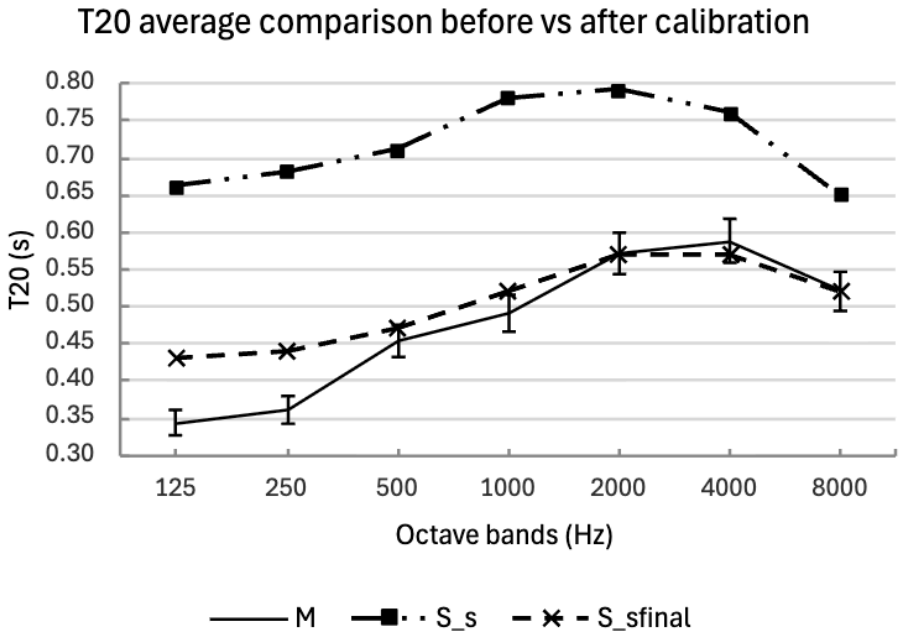


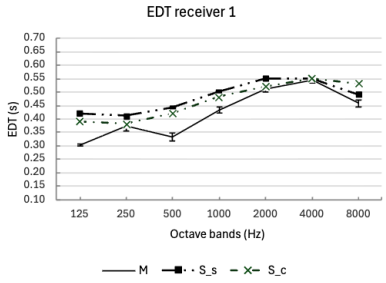
Figure 9.14 Comparison of octave-band T20 values between measurements (M), simulations (S_s) and the final simulation of the calibration (S_sfinal) . Error bars indicate the variability of the measured data

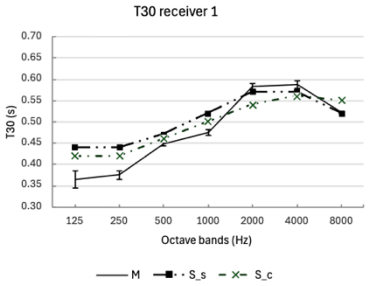
In light of the recent amendments, it is evident that the absorption and scattering coefficients embedded within the software now align with the values delineated in Table 9.4. Employing this revised array of acoustic parameters has yielded outcomes that exhibit a notable degree of concordance with the anticipated behaviour of the environment. In particular, with regard to the phenomenon of reverberation time, it was possible to achieve a configuration that satisfactorily reflected the response of the system across the entire frequency spectrum analysed. The ensuing trend is illustrated in Figure 9.14, which demonstrates how the modelling choices made have facilitated a more realistic calibration of the simulated acoustic field.

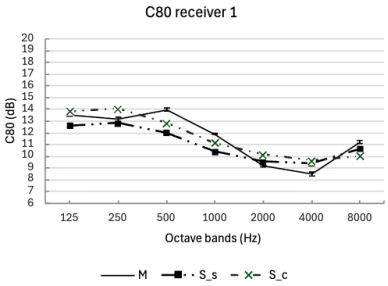
Table 9.4 Material properties adopted in the simulation model after calibration based on literature and manufacturer data.

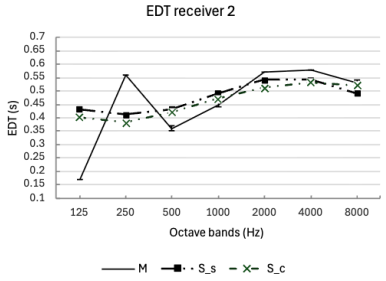
Component	Material	Scattering coefficient (μ_s)	Absorption coefficient (α)						
			125 Hz	200 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
walls	Plaster. gypsum. or lime. smooth finish on lath (Harris, 1991)[43]	0.3	0.185	0.135	0.085	0.06	0.05	0.038	0.025
Stone walls	Brick. Unglazed (Harris, 1991)[43]	0.3	0.033	0.033	0.036	0.048	0.055	0.077	0.099
ceiling	Knauf Danoline Plaza. sound-absorbing forated plasterboard (manufacturer data)[44]	0.05	0.583	0.790	0.905	0.780	0.698	0.563	0.438
floor	Wood parquet on counterfloor (Bobran. 1973)	0.3	0.22	0.157	0.105	0.1	0.05	0.1	0.15
windows	Double glazing. 2-3mm glass. >30mm gap (Fasold Winkler, 1976)[45]	0.05	0.165	0.1	0.06	0.044	0.022	0.022	0.022
studio window	Double glazing. 2-3mm glass. >30mm gap (Fasold Winkler, 1976)[45]	0.05	0.165	0.1	0.06	0.044	0.022	0.022	0.022
steel doors	Solid timber door (akustik.ua)[43,46]	0.05	0.11	0.066	0.088	0.11	0.11	0.11	0.11
bench	Wood parquet in asphalt on concrete (Harris. 1991)[43]	0.05	0.053	0.053	0.093	0.079	0.073	0.093	0.101
cylindrical diffusing elements	Thin Plywood paneling (Ref. Dalenback. CATT)	0.89 *	0.562	0.331	0.21	0.188	0.166	0.066	0.06
Ceiling suspended diffusive panel	8mm wooden ceiling suspended panel (Cox & D'Antonio, 2003)[47]	0.05	0.055	0.088	0.088	0.11	0.11	0.11	0.05
Schroeder diffuser	1D N=7 QRD. design freq.= 500 Hz (Cox & D'Antonio)	0.87 *	0.11	0.1	0.07	0.08	0.06	0.06	0.06
Bass traps	pressed polystyrene EPS (Cox & D'Antonio, 2004)	0.05	0.2	0.1	0.05	0.022	0.039	0.066	0.077

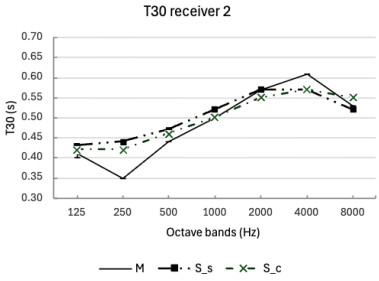
D. Receiver-wise breakdown of S_c, S_s and M on Odeon

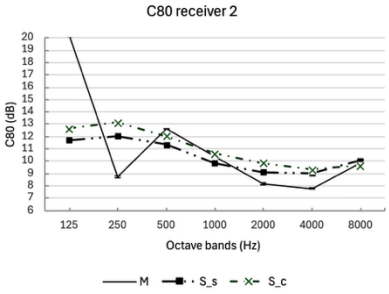


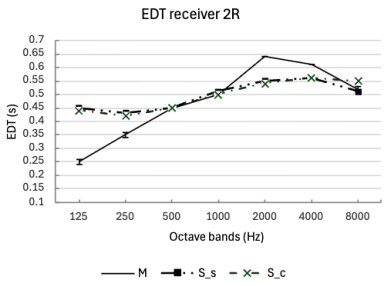


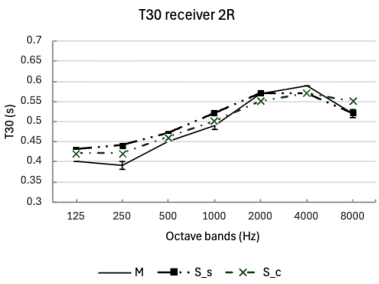


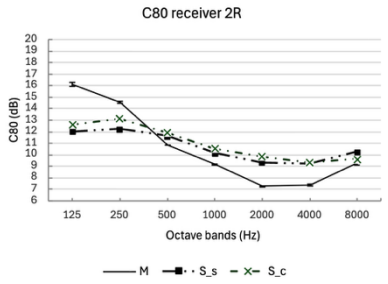


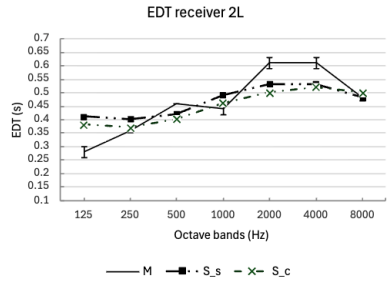


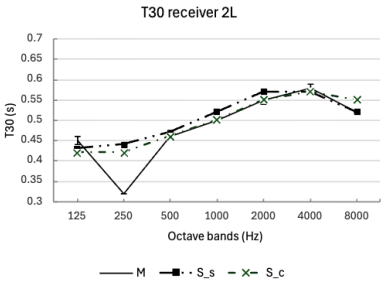


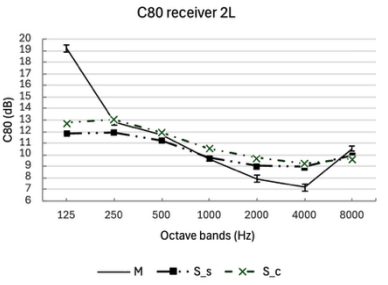


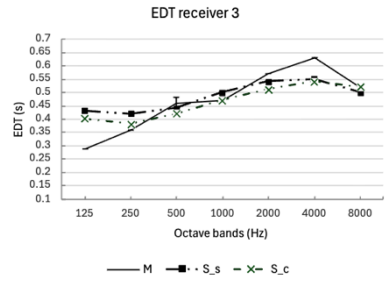


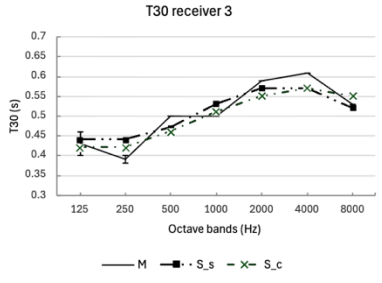


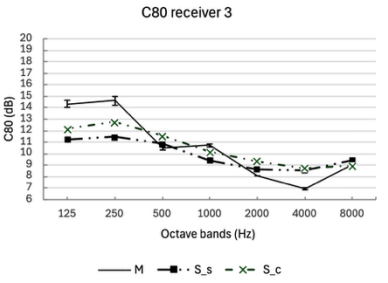


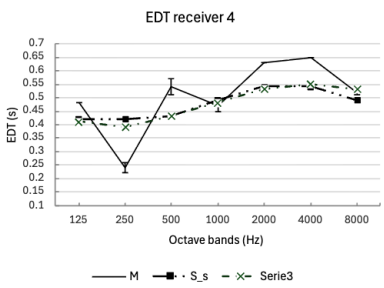


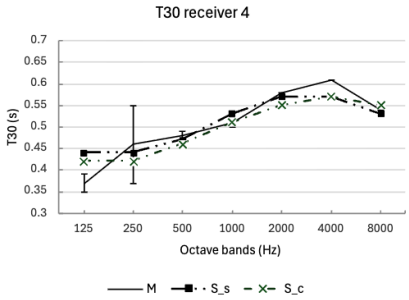


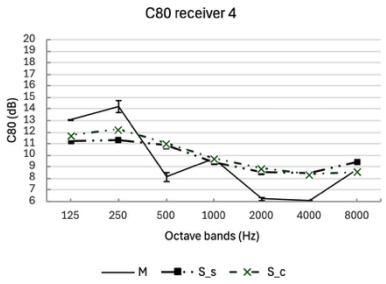







EDT receiver 4
EDT (s)
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
125
250
500
1000
2000
4000
8000
Octave bands (Hz)
M
S_s
Serie3


T30 receiver 4
T30 (s)
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
125
250
500
1000
2000
4000
8000
Octave bands (Hz)
M
S_s
S_c


C80 receiver 4
C80 (dB)
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
125
250
500
1000
2000
4000
8000
Octave bands (Hz)
M
S_s
S_c

E. Receiver-wise breakdown of room configurations

In this appendix, the acoustic parameters are reported and compared on a receiver-by-receiver basis. While the main body of the work presents spatially averaged values in order to highlight global trends, the following graphs allow a detailed inspection of the local acoustic behaviour at each receiver position. This representation provides a deeper understanding of spatial variability and helps assess the robustness of the conclusions drawn from the averaged results.
Only EDT, T_{30} and C_{80} are reported, as they capture the most relevant variations in both reverberation behaviour and sound clarity, while the remaining parameters exhibited similar trends.
In Figure 9.15 it is depicted the behaviour of the different configurations with the respect to the complex model in in Figure 9.16. the comparison with the respect to the simplified model with the bidimensional version of the cilindrical panel and the Schroeder diffuser.

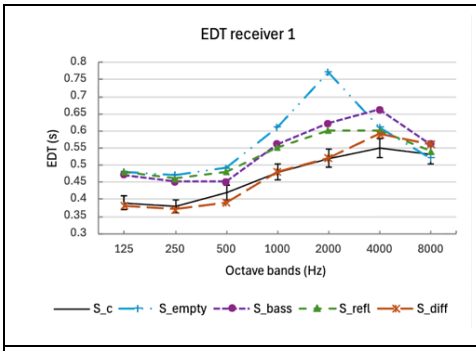
One thing that is now more visible locally, making attention to the values of T_{30} , than in the graphs showing the average values is the small peak for empty room between 250 Hz and 1 kHz. This indicates that, in the empty room configuration, the sound energy reaching position one exhibits a slightly longer decay time. This could be caused by the fact that position one is the nearest to the source.

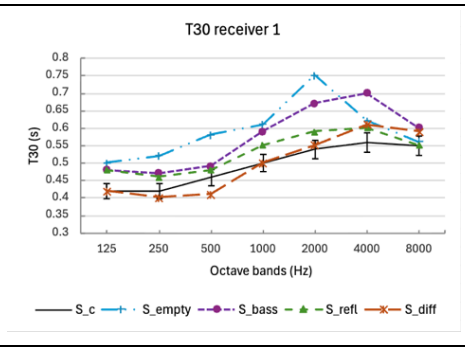
In particular, the addition of the cylindrical panels and the Schroeder diffuser is highly effective in this frequency range when the room is arranged as shown, as well as in the 2000 Hz range that we analysed previously.

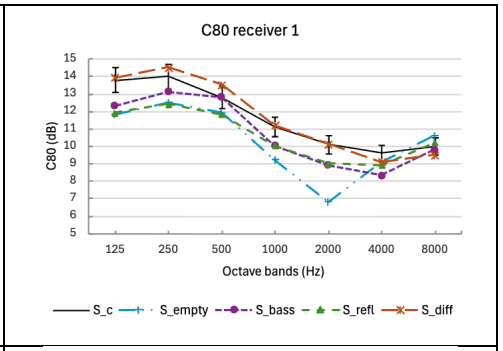
For these frequencies and in this position, the phenomenon has no significant effect on the clarity parameters or the early decay time. A slight influence on the reverberation time can be observed in positions 2, while in the remaining positions it is essentially negligible.

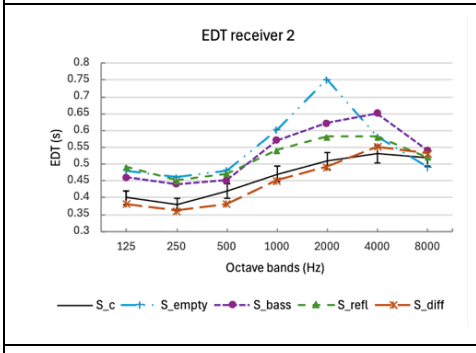
Analysis of the C_{80} parameter graphs reveals that almost all room configurations, with the exception of the empty configuration, are substantially consistent with the complex reference model at frequencies close to 8000 Hz, for all positions considered. This behaviour is not evident in comparisons with the simplified model, which, as discussed in previous sections, exhibits less acoustic stability than the model with three-dimensional diffusers. The same phenomenon can be clearly observed in the EDT plots.

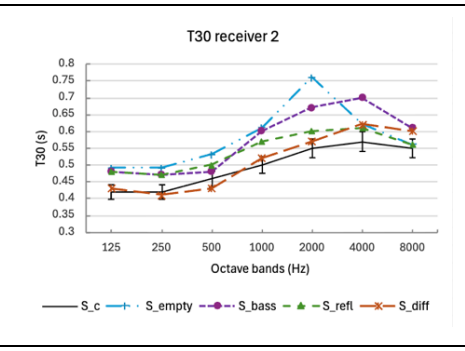
Another interesting observation is the positive imbalance in mid-high frequencies, accentuated for receivers further from the source (2L, 3 and 4), especially with regard to EDT and T_{30}. This phenomenon is much more manageable in the 3D model (S_s) than in the simplified model (S_c). In fact, the configuration with reflective panels suspended from the ceiling contains this behaviour as well as the configuration with cylindrical panels and the Schroeder diffuser. This configuration comes very close to the reverberation time of the complete S_c model, and is much more effective than the configuration with two-dimensional diffuser panels, both compared to the reflective ceiling panels and, above all, compared to the three-dimensional diffusers.

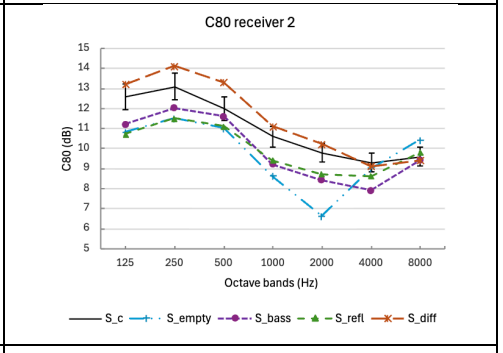

EDT receiver 1
EDT (s)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff

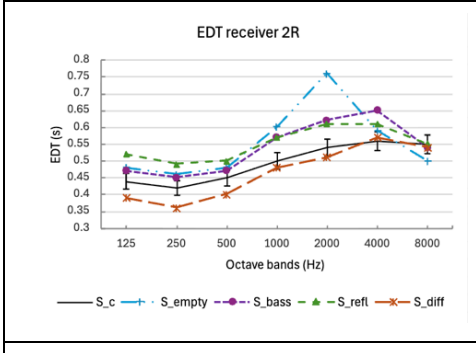

T30 receiver 1
T30 (s)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff

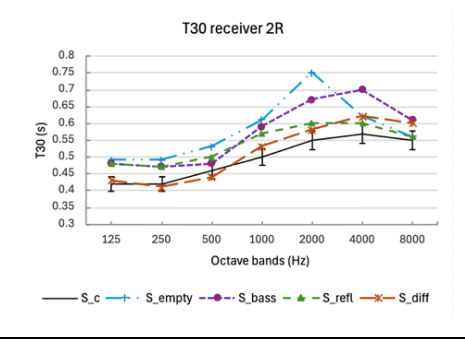

C80 receiver 1
C80 (dB)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff

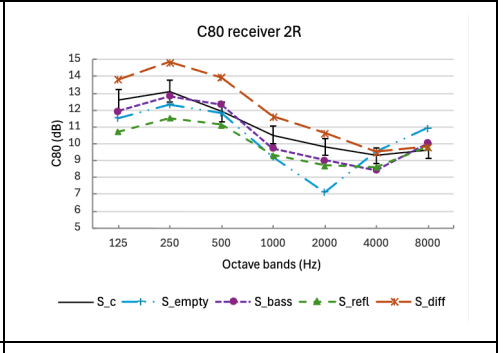

EDT receiver 2
EDT (s)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff

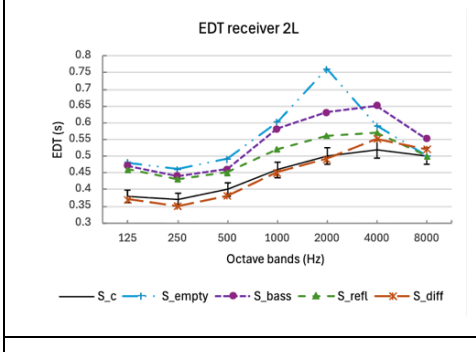

T30 receiver 2
T30 (s)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff

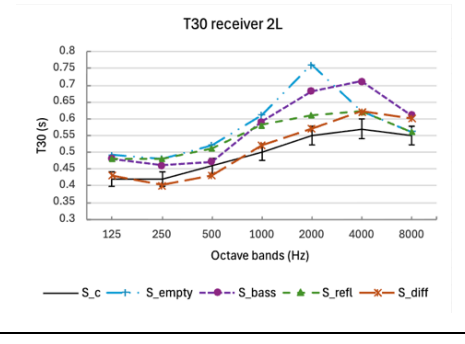

C80 receiver 2
C80 (dB)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff

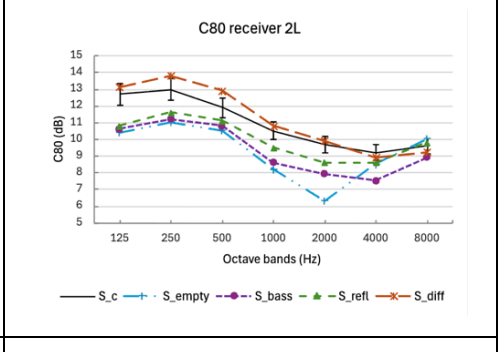

EDT receiver 2R
EDT (s)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff

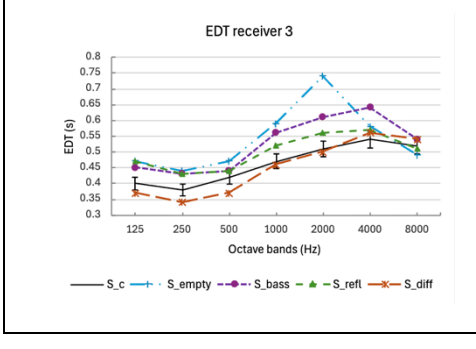

T30 receiver 2R
T30 (s)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff

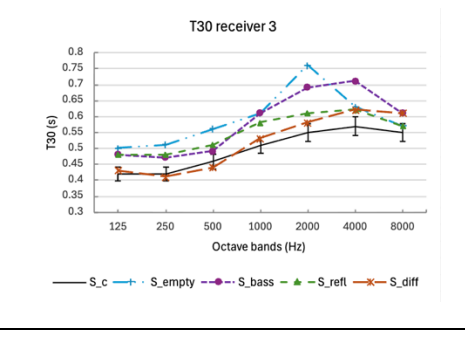

C80 receiver 2R
C80 (dB)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff

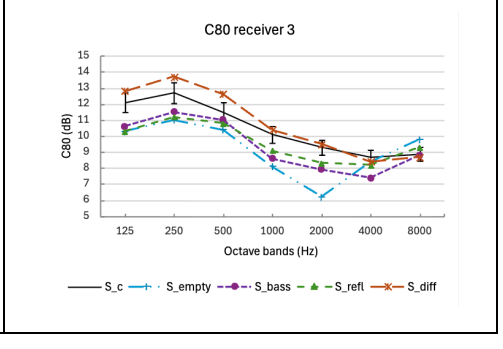

EDT receiver 2L
EDT (s)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff


T30 receiver 2L
T30 (s)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff


C80 receiver 2L
C80 (dB)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff


EDT receiver 3
EDT (s)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff


T30 receiver 3
T30 (s)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff


C80 receiver 3
C80 (dB)
Octave bands (Hz)
S_c
S_empty
S_bass
S_refl
S_diff

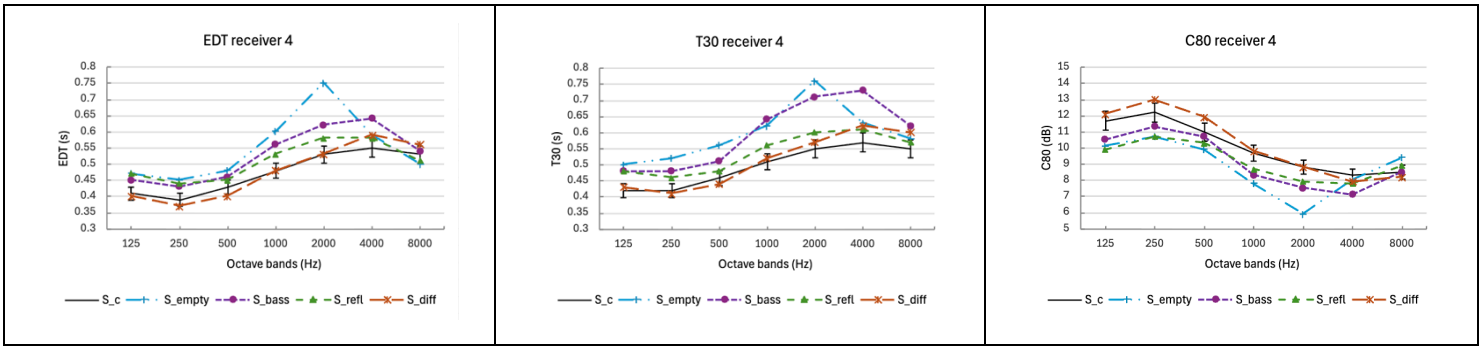
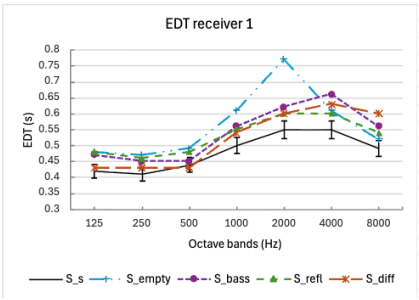
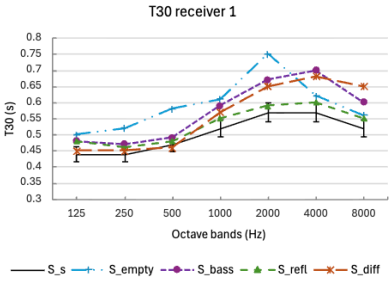
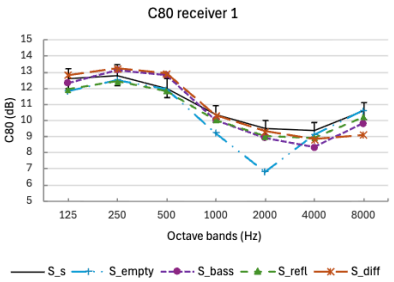
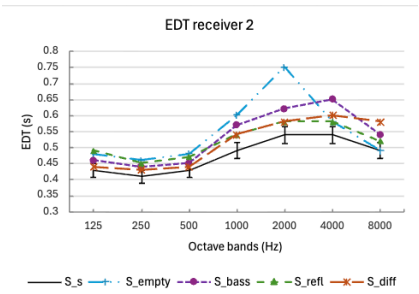


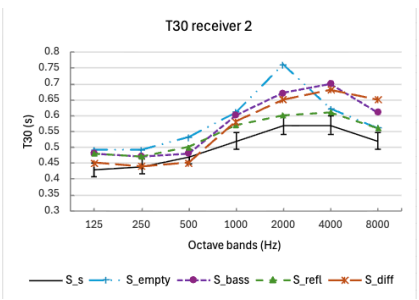
Figure 9.15 Key acoustic parameters of different configuration of the room with the respect to the complex simulated model

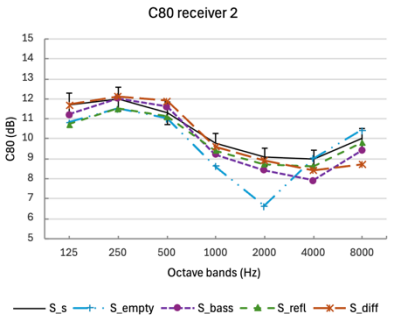

EDT receiver 1
EDT (s)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff

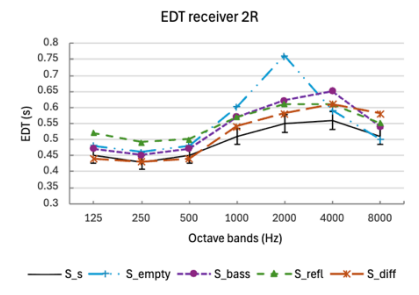

T30 receiver 1
T30 (s)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff

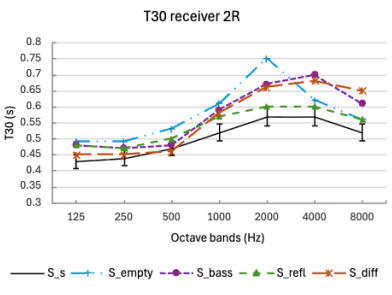

C80 receiver 1
C80 (dB)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff

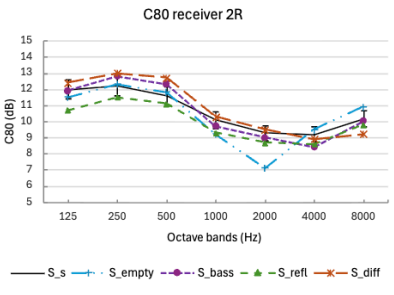

EDT receiver 2
EDT (s)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff

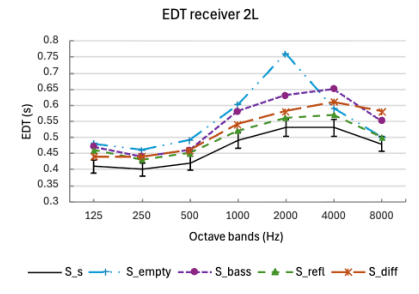

T30 receiver 2
T30 (s)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff

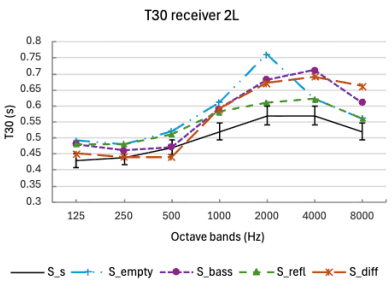

C80 receiver 2
C80 (dB)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff

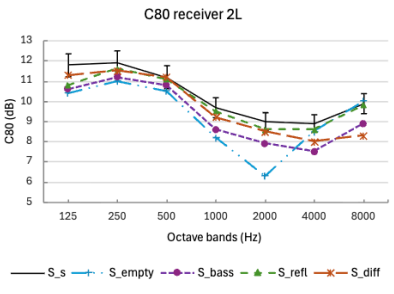

EDT receiver 2R
EDT (s)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff

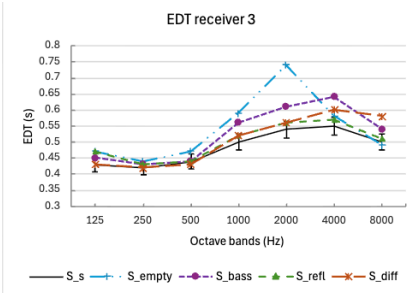

T30 receiver 2R
T30 (s)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff

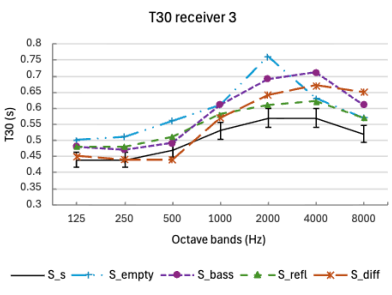

C80 receiver 2R
C80 (dB)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff

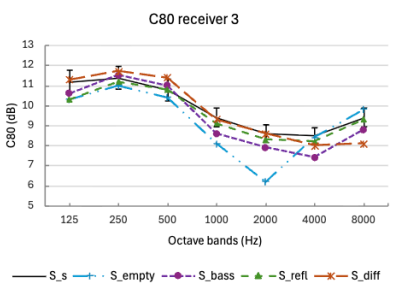

EDT receiver 2L
EDT (s)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff


T30 receiver 2L
T30 (s)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff


C80 receiver 2L
C80 (dB)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff


EDT receiver 3
EDT (s)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff


T30 receiver 3
T30 (s)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff


C80 receiver 3
C80 (dB)
Octave bands (Hz)
S_s
S_empty
S_bass
S_refl
S_diff

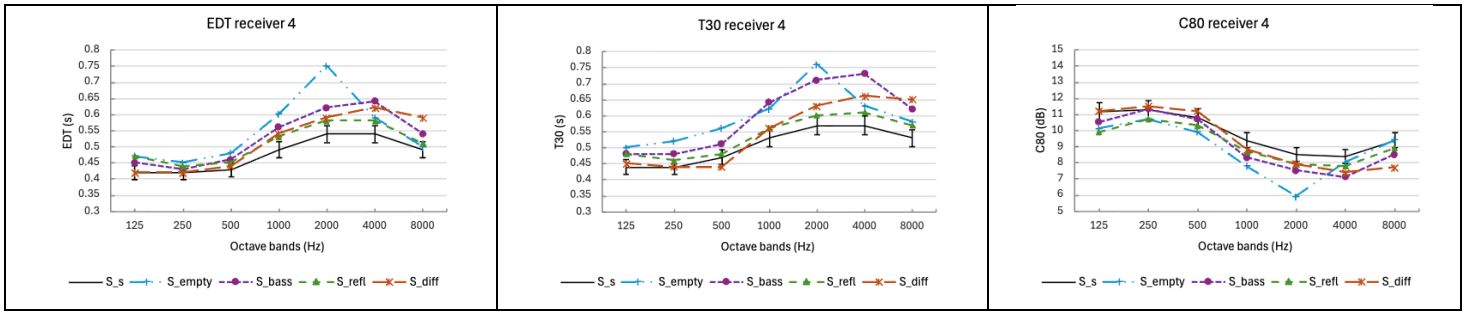


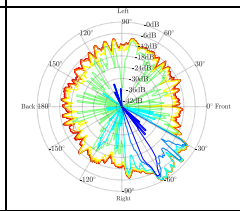
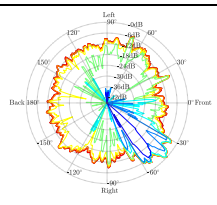
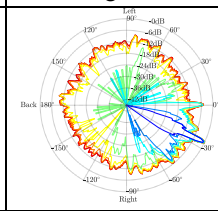
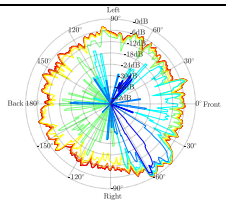
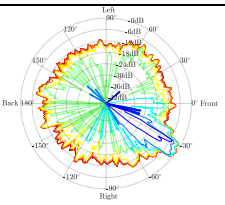
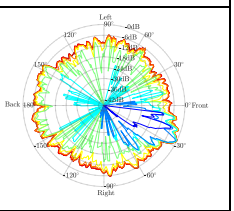
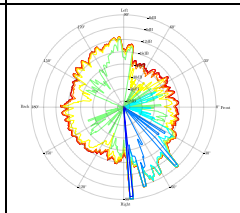
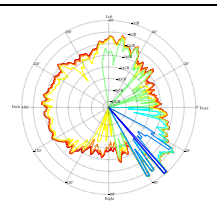
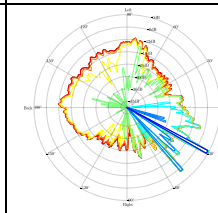
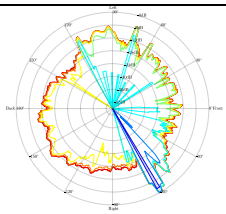
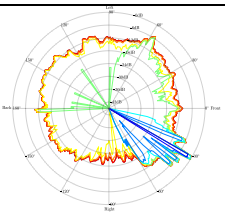
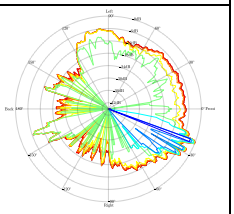
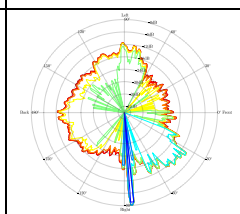
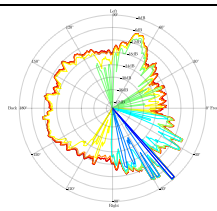
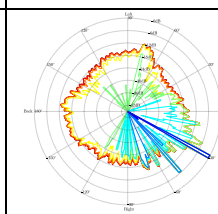
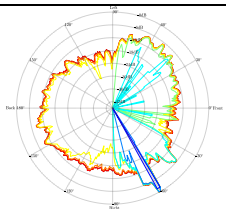
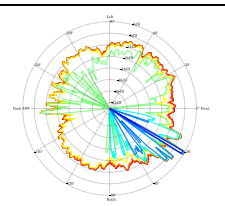
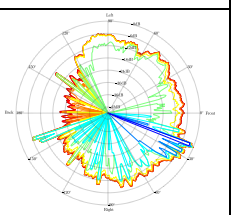
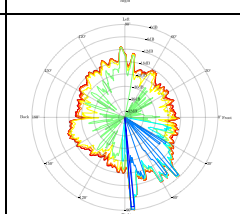
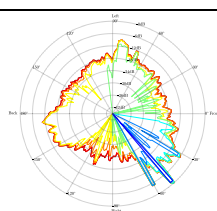
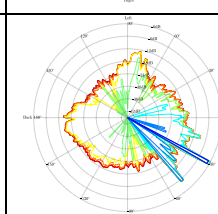
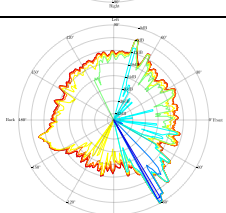
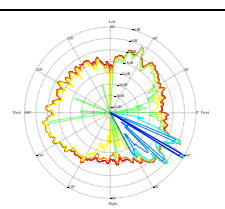
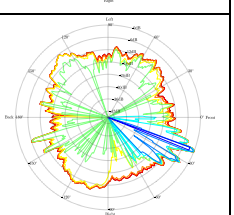
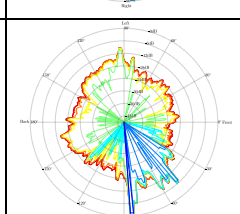
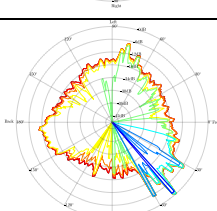
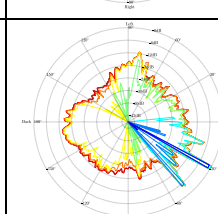
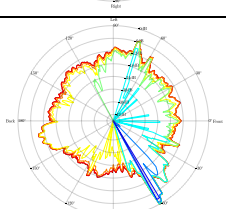
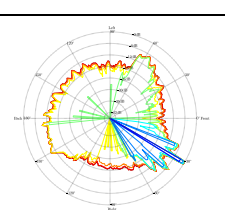
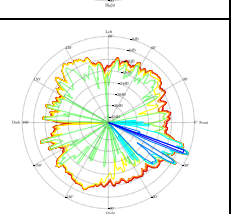
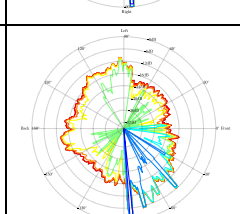
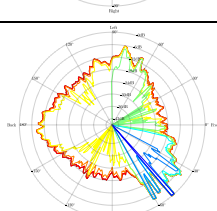
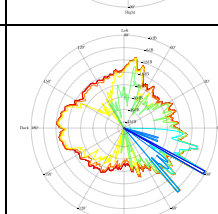
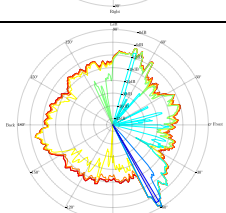
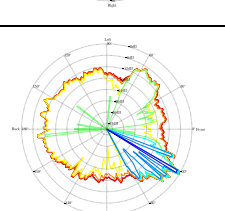
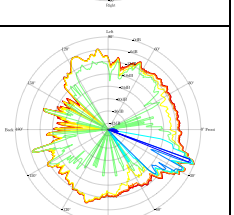
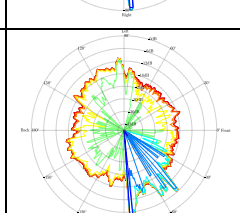
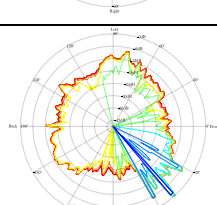
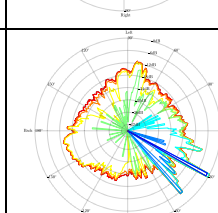
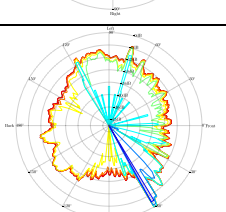
Figure 9.16 key acoustic parameters of the different room configuration with the respect to the simplified simulated model

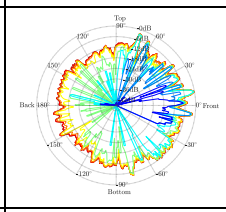
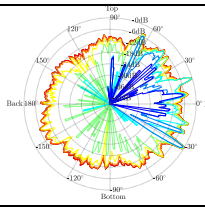
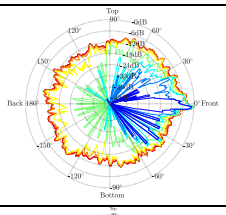
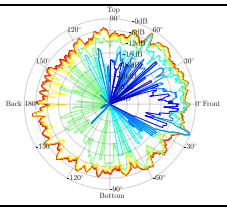
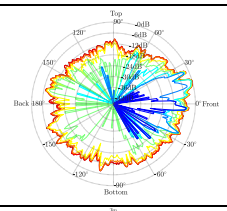
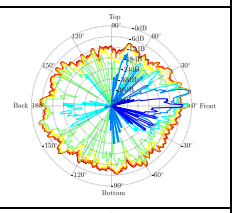
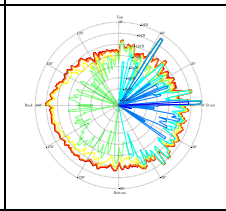
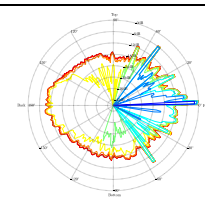
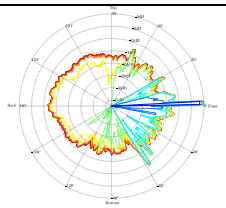
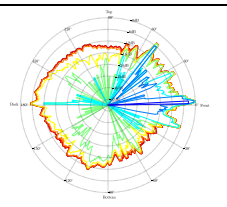
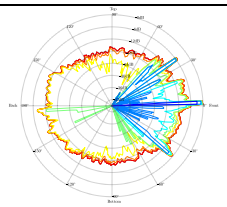
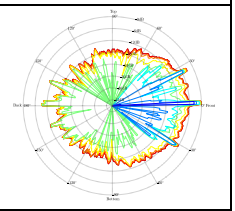
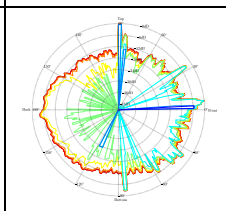
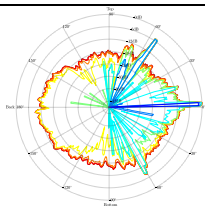
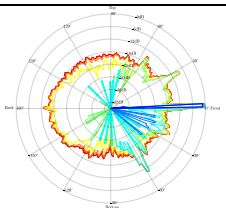
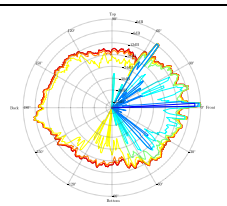
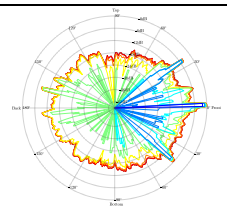
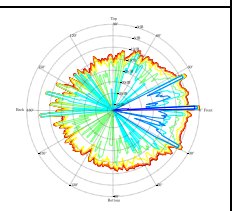
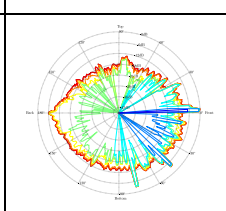
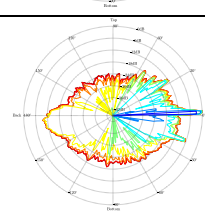
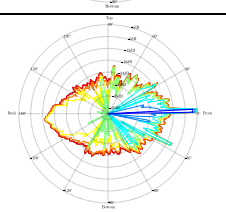
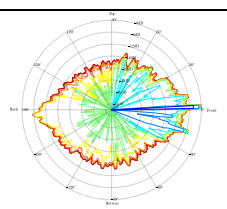
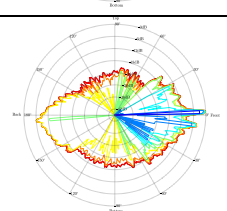
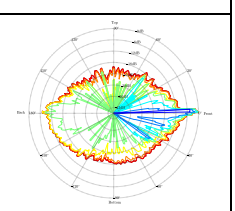
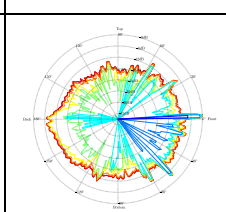
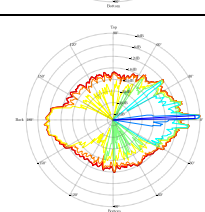
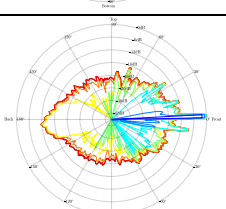
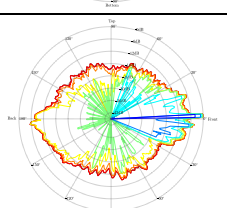
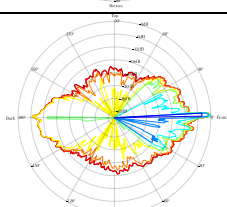
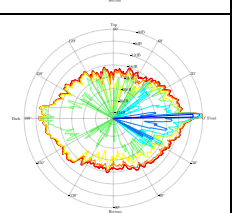
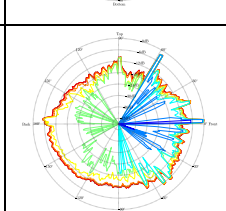
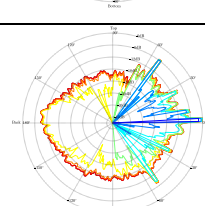
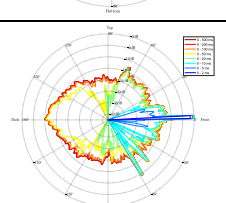
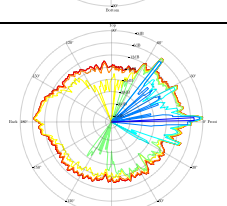
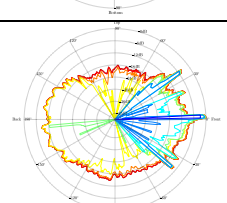
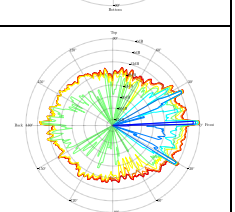
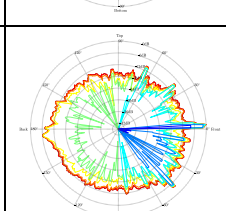
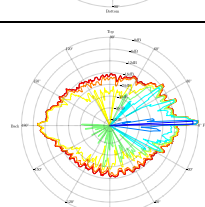
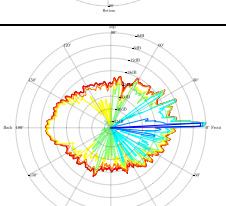
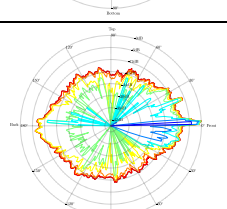
F. Receiver-wise decay roses

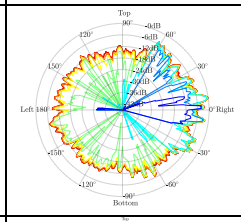
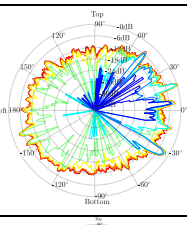
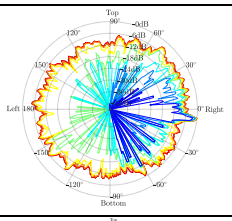
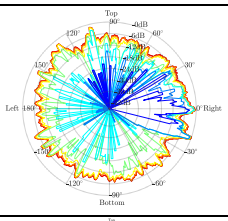
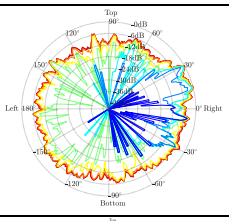
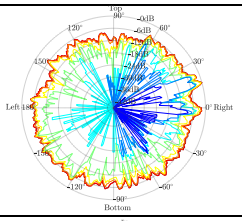
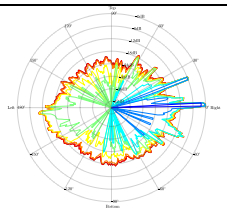
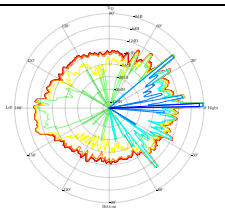
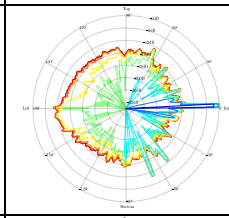
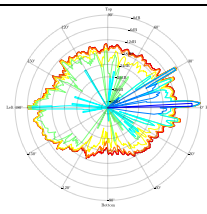
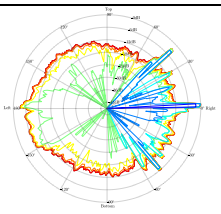
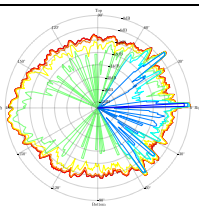
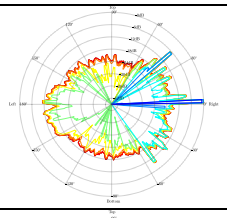
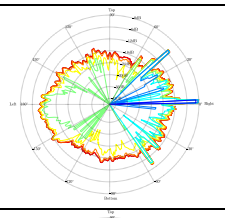
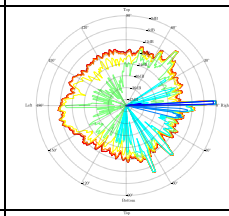
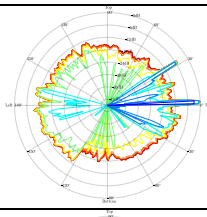
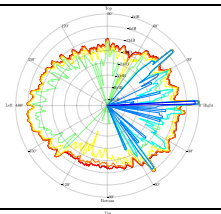
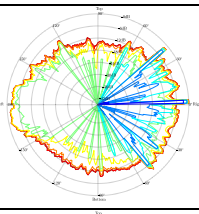
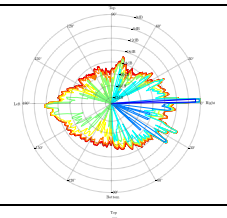
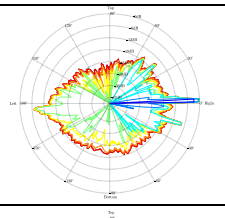
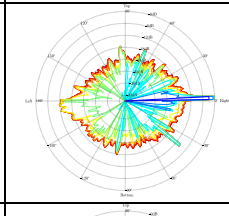
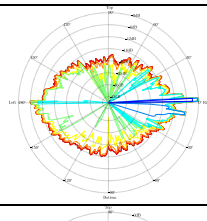
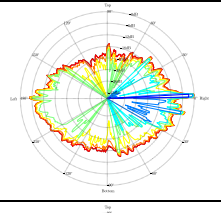
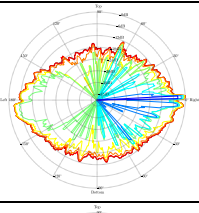
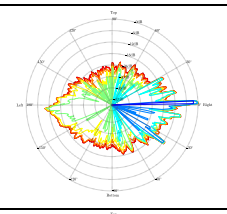
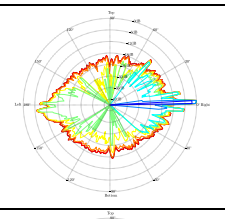
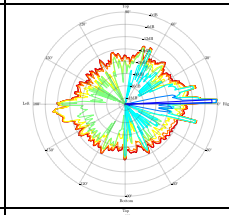
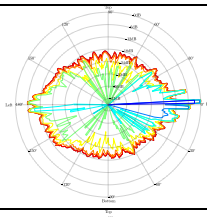
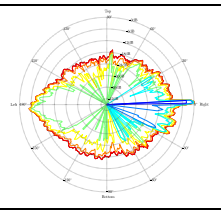
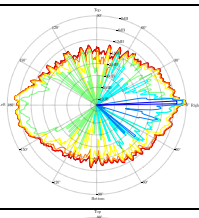
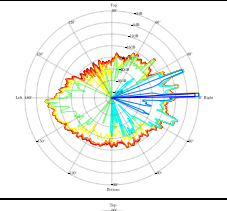
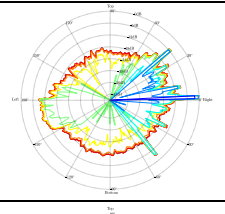
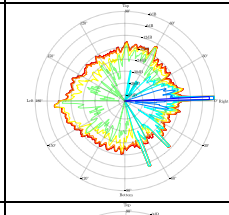
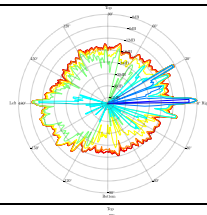
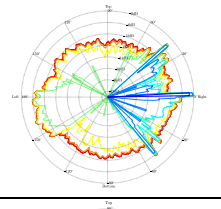
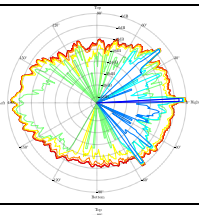
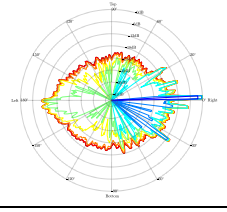
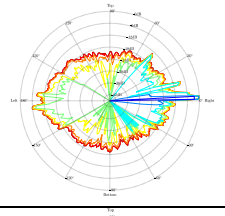
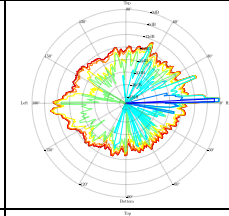
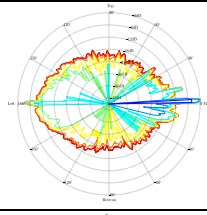
Rose decay plots have been derived from the ambisonic impulse responses for the acoustic configurations that have been measured and simulated.

The three tables represent different spatial sections of the sound field: the first table shows the lateral sections, the second table the median sections, and the third table the transverse sections.

Within each table, the rows correspond to the various acoustic configurations that were considered in the study. These are as follows: measured condition (M), simplified Odeon model with 2D diffusers (S_s), detailed Odeon model with 3D diffusers (S_c), empty-room simulation (S_{empty}), simulation including only bass traps (S_{bass}), simulation including only reflective ceiling panels (S_{refl}), simulation including only 3D diffusers (S_{diff3D}), and simulation including only 2D diffusers (S_{diff2D}). The columns are aligned with the receiver positions (Receiver 1, Receiver 2, Receiver 2 right, Receiver 2 left, Receiver 3 and Receiver 4). Each polar plot is representative of the directional distribution of the energy decay extracted from the Ambisonic impulse responses, thus facilitating a qualitative comparison of the effects that different acoustic treatments have on the spatial redistribution of reflected sound ener

	R1 lateral	R2 lateral	R2 right lateral	R2 left lateral	R3 lateral	R4 lateral
M						
S_c						
S_s						
S_{empty}						
S_{bass}						
S_{refl}						
S_{diff3D}						
S_{diff2D}						

	R1 median	R2 median	R2 right median	R2 left median	R3 median	R4 median
M						
S_c						
S_s						
S_{empty}						
S_{bass}						
S_{refl}						
S_{diff3D}						
S_{diff2D}						

	R1 trasnverse	R2 transverse	R2 right trasnverse	R2 left trasnverse	R3 trasnverse	R4 trasnverse
M						
S_c						
S_s						
S_{empty}						
S_{bass}						
S_{refl}						
S_{diff3D}						
S_{diff2D}						

G. Frequency response plot S_{empty} and S_c

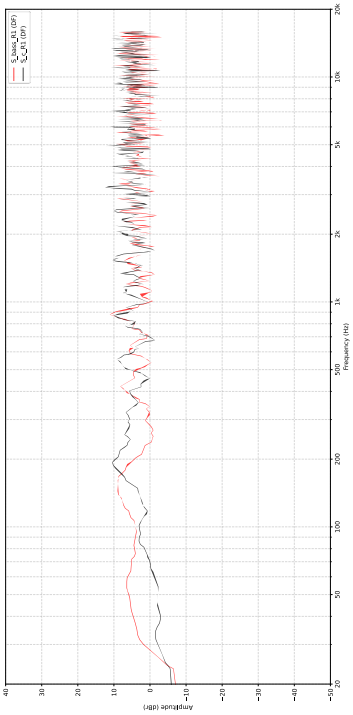


Figure 9.17 S_{empty} and S_c from receiver 1

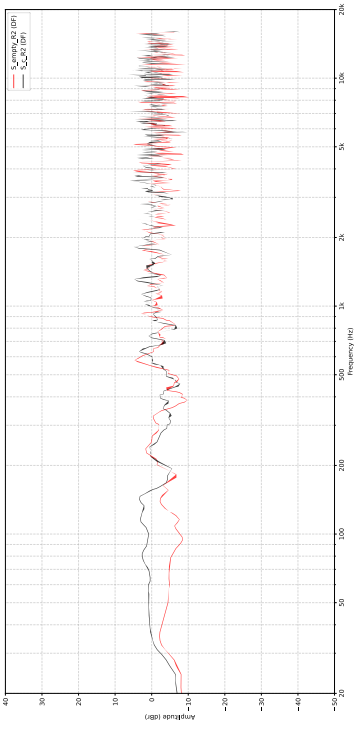


Figure 9.18 S_{empty} and S_c from receiver 2

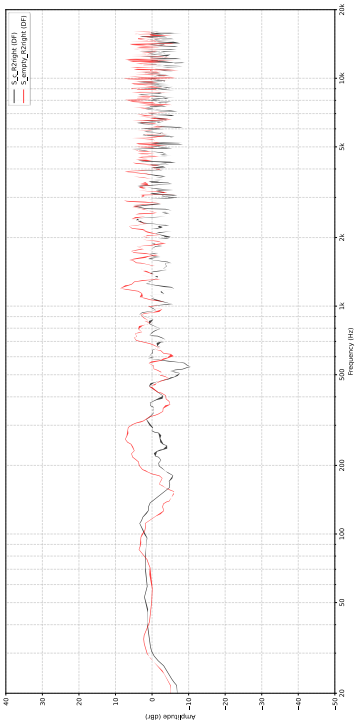


Figure 9.19 S_{empty} and S_c from receiver 2 right

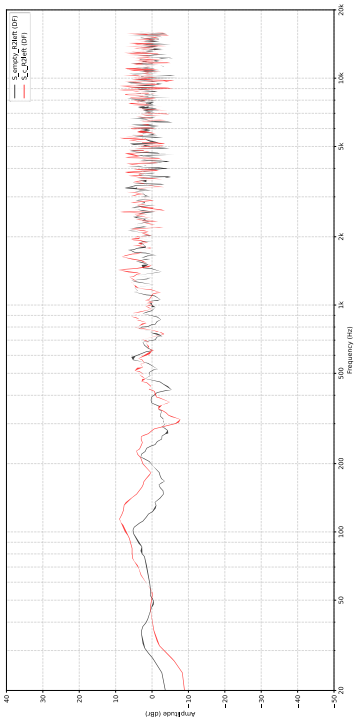


Figure 9.20 S_{empty} and S_c from receiver 2 left

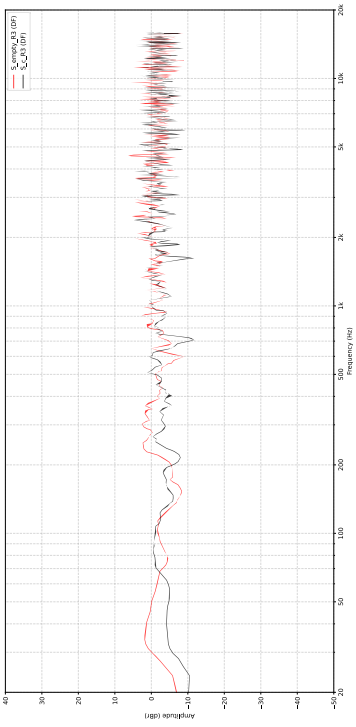


Figure 9.21 S_{empty} and S_c from receiver 3

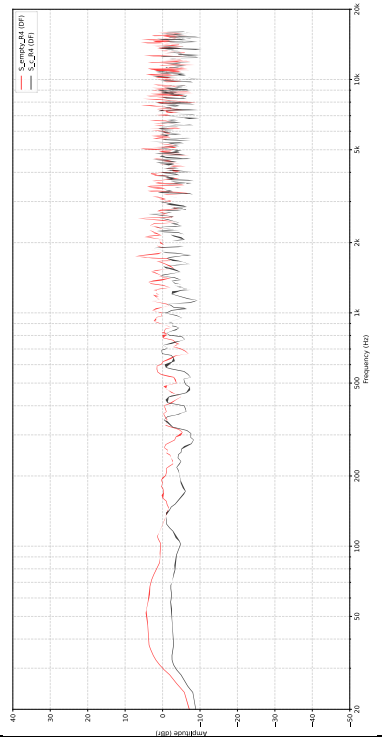


Figure 9.22 S_{empty} and S_c from receiver 4

H. Frequency response plot S_{bass} and S_c

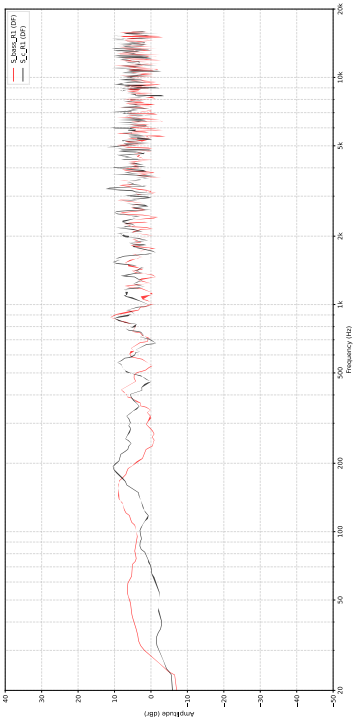


Figure 9.23 S_{bass} and S_c from receiver 1

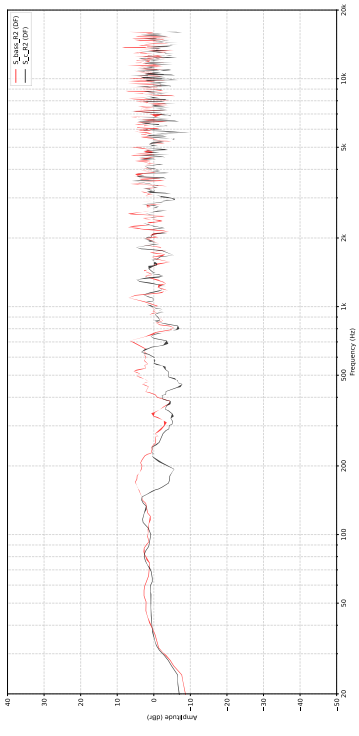


Figure 9.24 S_{bass} and S_c from receiver 2

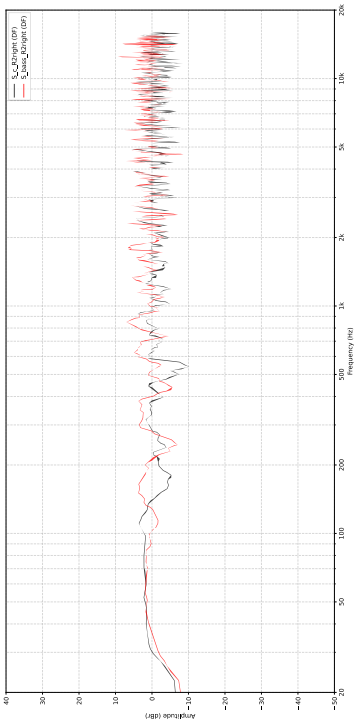


Figure 9. S_{bass} and S_c from receiver 2 right

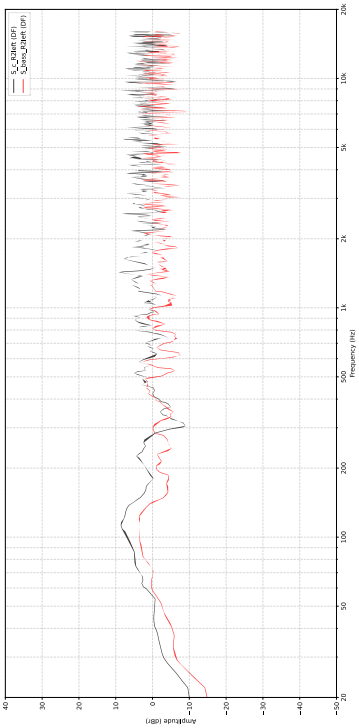


Figure 9.25 S_{bass} and S_c from receiver 2 left

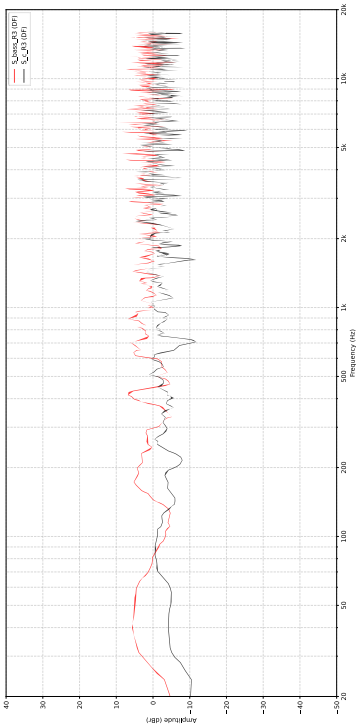


Figure 9.26 S_{bass} and S_c from receiver 3

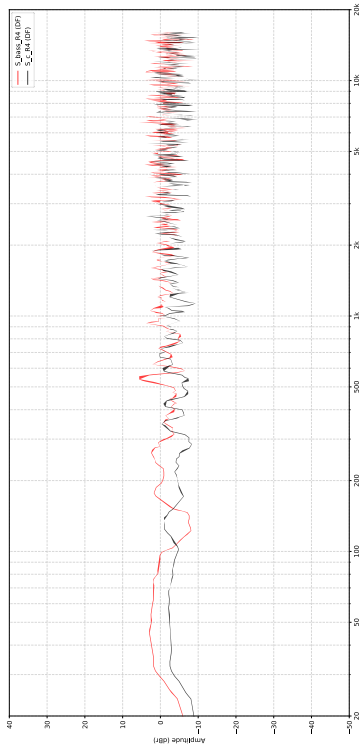


Figure 9.27 S_{bass} and S_c from receiver 4

I. Frequency response plot S_{refl} and S_c

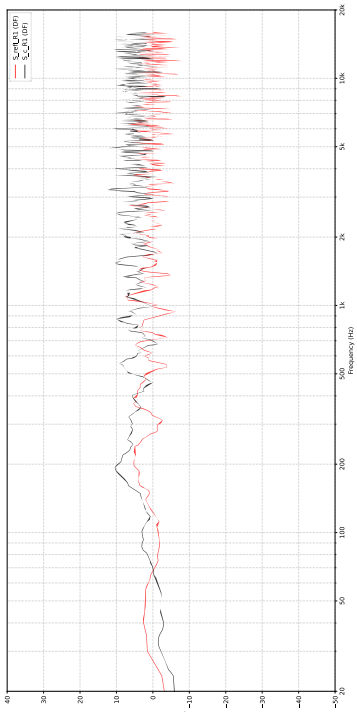


Figure 9.28 S_{refl} and S_c from receiver 1

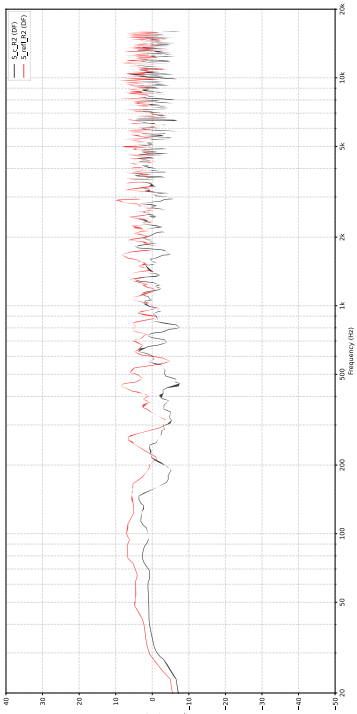


Figure 9.29 S_{refl} and S_c from receiver 2

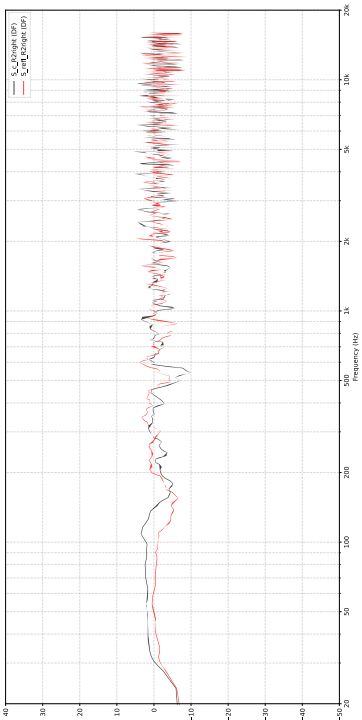


Figure 9.30 S_{refl} and S_c from receiver 2 right

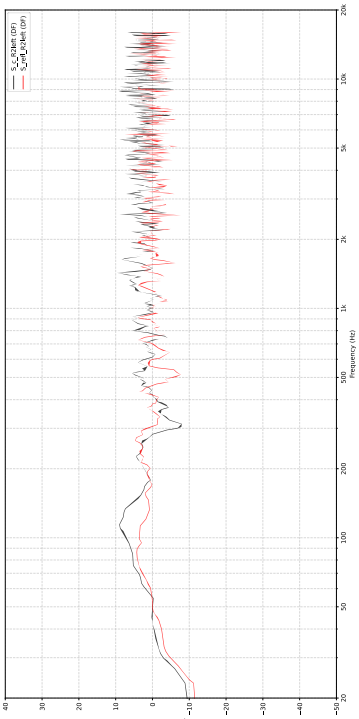


Figure 9.31 S_{refl} and S_c from receiver 2 left

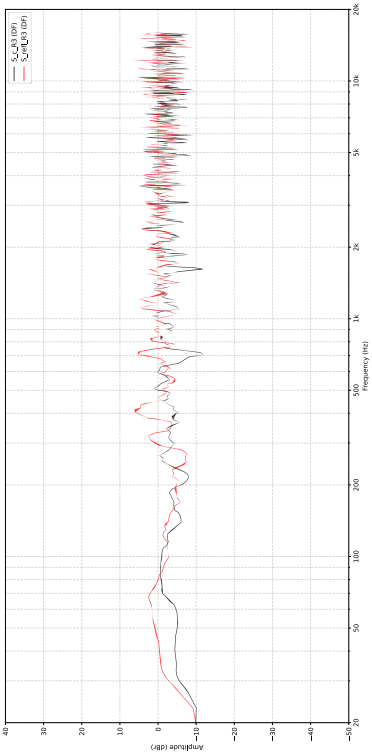


Figure 9.32 S_{refl} and S_c from receiver 3

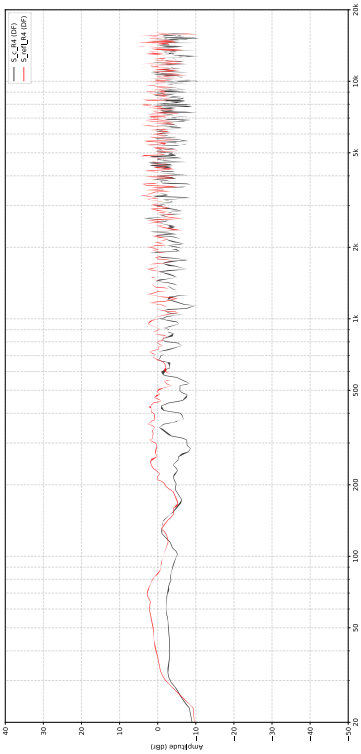


Figure 9.33 S_{refl} and S_c from receiver 4

J. Frequency response plot S_{diff3D} and S_c

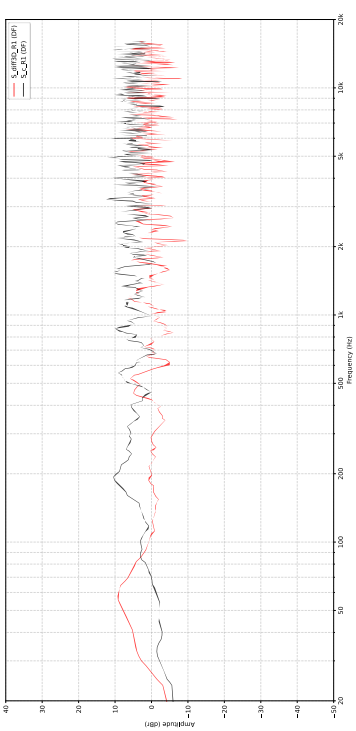


Figure 9.34 S_{diff3D} and S_c from receiver 1

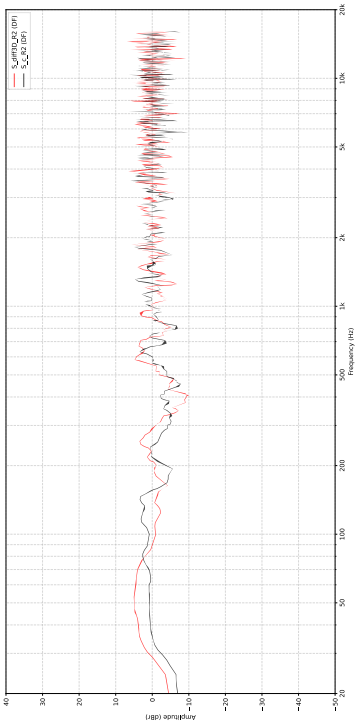


Figure 9.35 S_{diff3D} and S_c from receiver 2

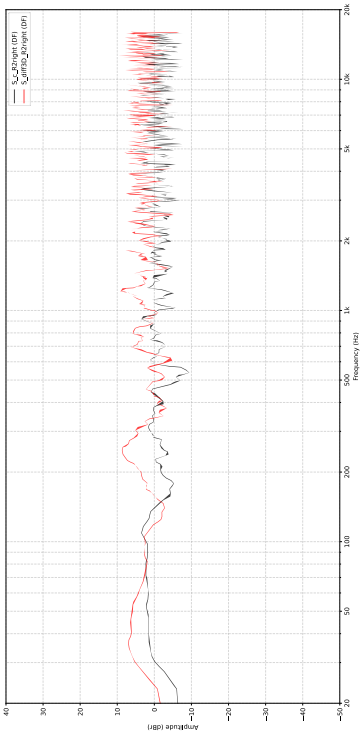


Figure 9.36 S_{diff3D} and S_c from receiver 2 right

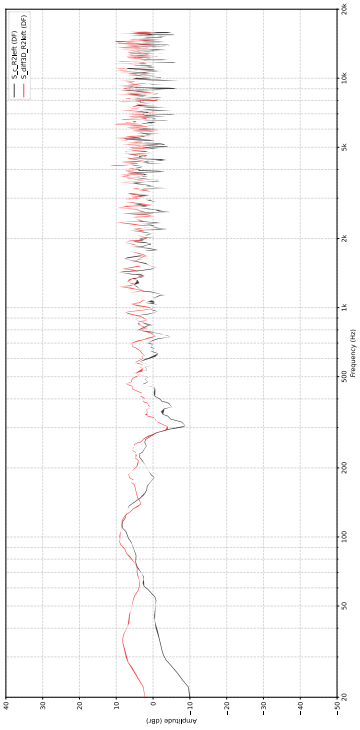


Figure 9.37 S_{diff3D} and S_c from receiver 2 left

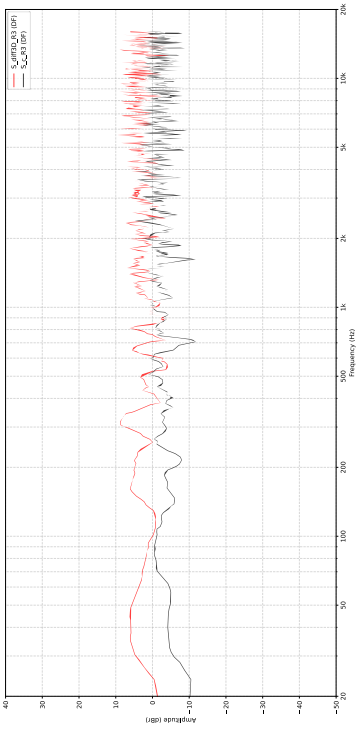


Figure 9.38 S_{diff3D} and S_c from receiver 3

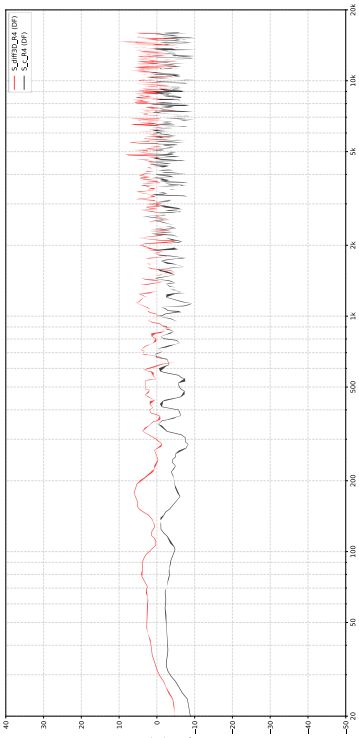


Figure 9.39 S_{diff3D} and S_c from receiver 4

K. Frequency response plot S_{empty} and S_s

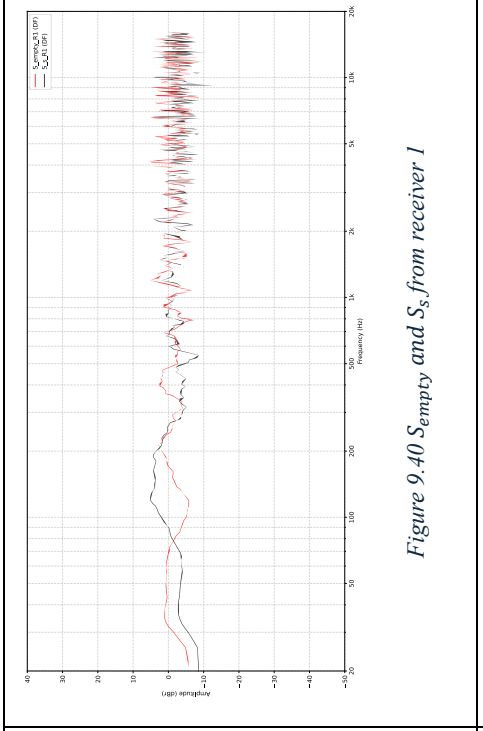


Figure 9.40 S_{empty} and S_s from receiver 1

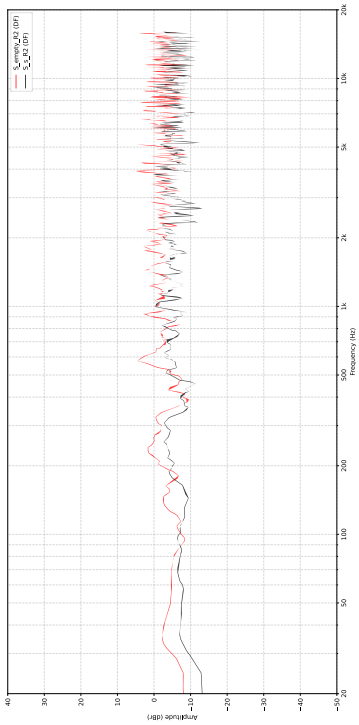


Figure 9.41 S_{empty} and S_s from receiver 2

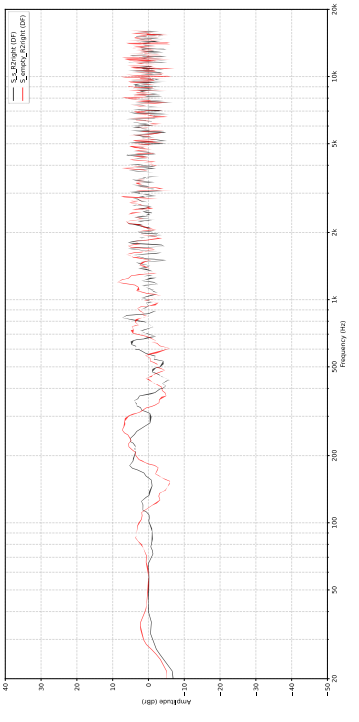


Figure 9.42 S_{empty} and S_s from receiver 2 right

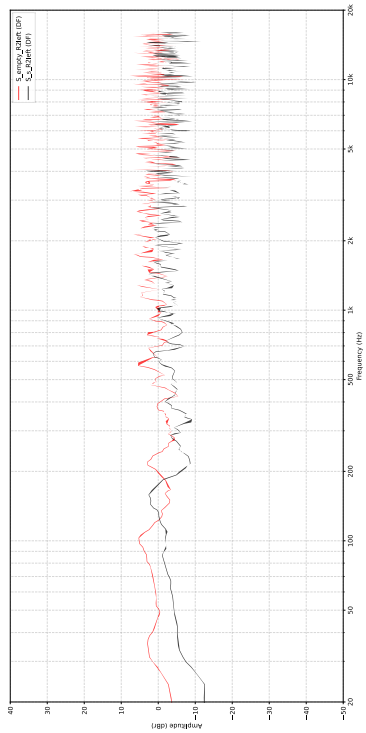


Figure 9.43 S_{empty} and S_s from receiver 2 left

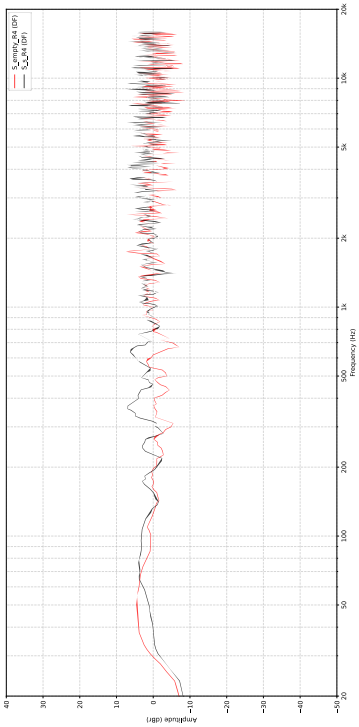


Figure 9.44 S_{empty} and S_s from receiver 3

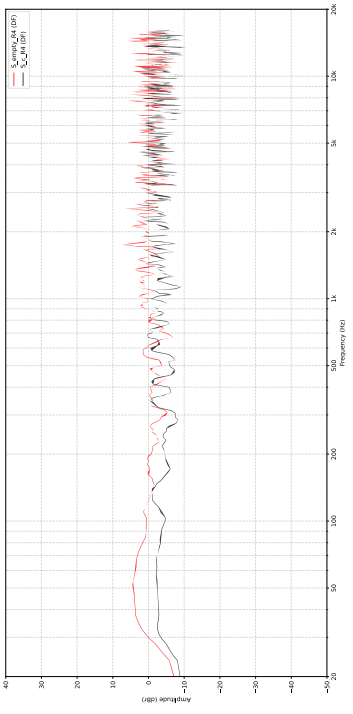


Figure 9.45 S_{empty} and S_s from receiver 4

L. Frequency response plot S_{bass} and S_s

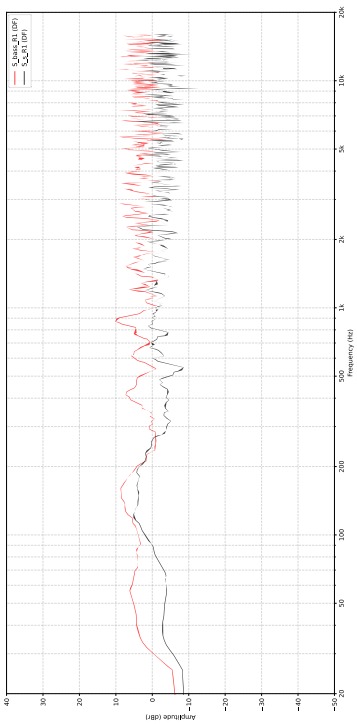


Figure 9.46 S_{bass} and S_s from receiver 1

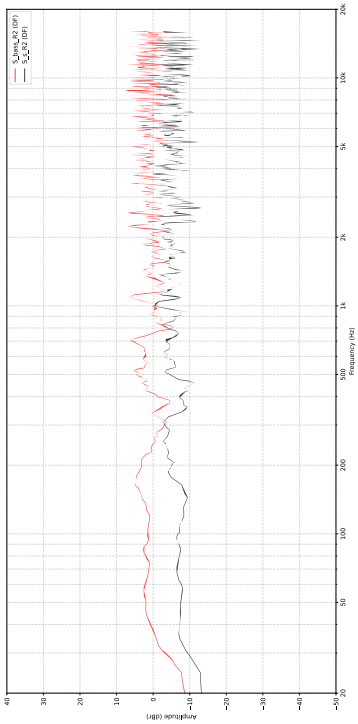


Figure 9.47 S_{bass} and S_s from receiver 2

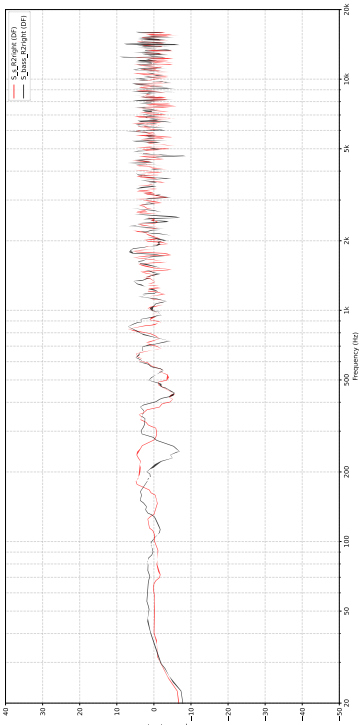


Figure 9.48 S_{bass} and S_s from receiver 2 right

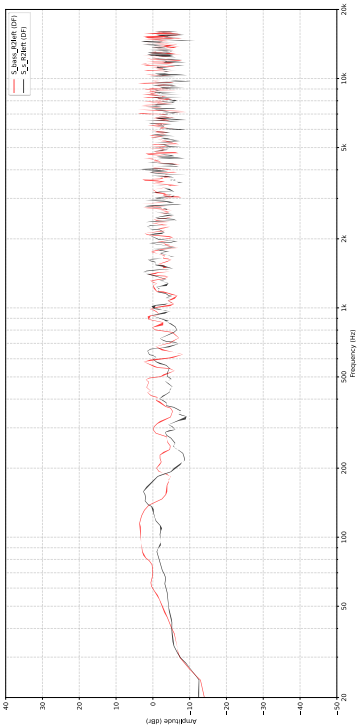


Figure 9.49 S_{bass} and S_s from receiver 2 left

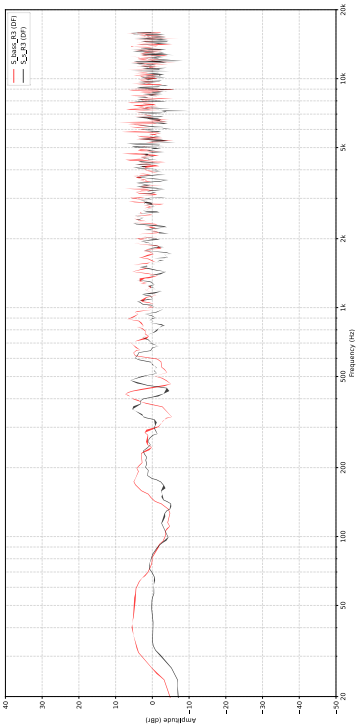


Figure 9.50 S_{bass} and S_s from receiver 3

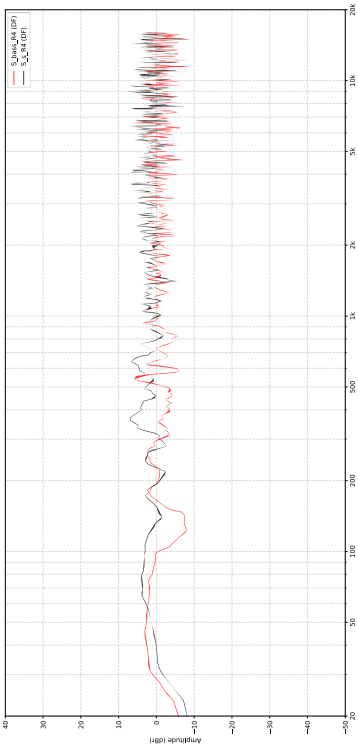


Figure 9.51 S_{bass} and S_s from receiver 4

M. Frequency response plot S_{refl} and S_s

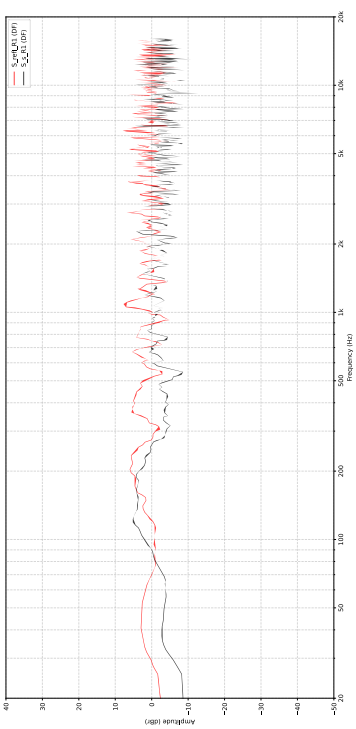


Figure 9.52 S_{refl} and S_s from receiver 1

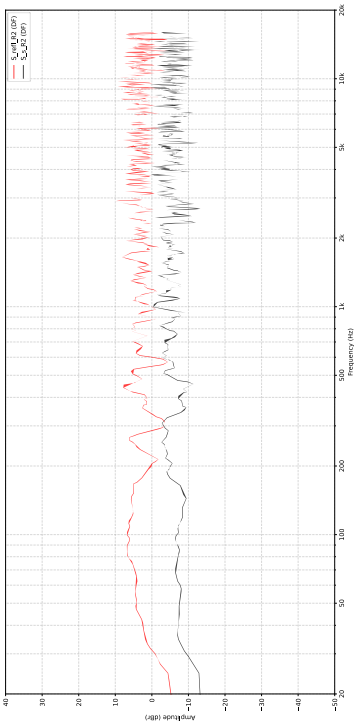


Figure 9.53 S_{refl} and S_s from receiver 2

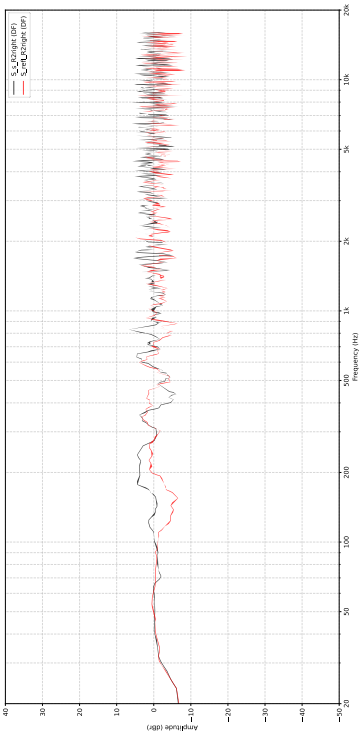


Figure 9.54 S_{refl} and S_s from receiver 2 right

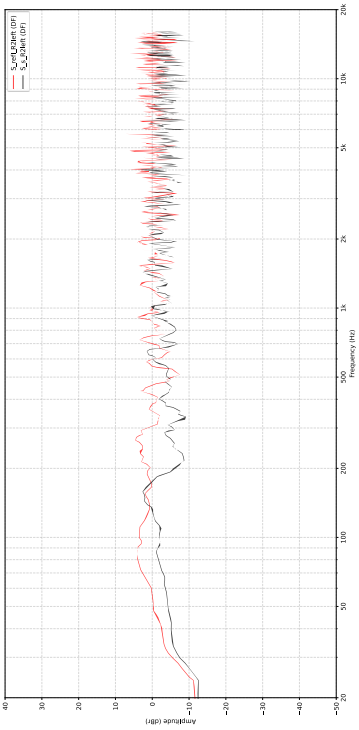


Figure 9.55 S_{refl} and S_s from receiver 2 left

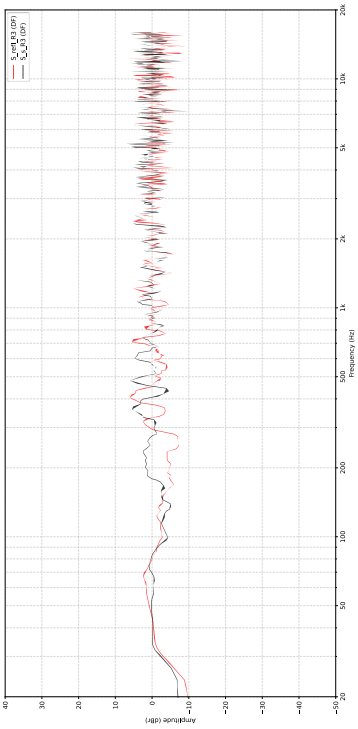


Figure 9.56 S_{refl} and S_s from receiver 3

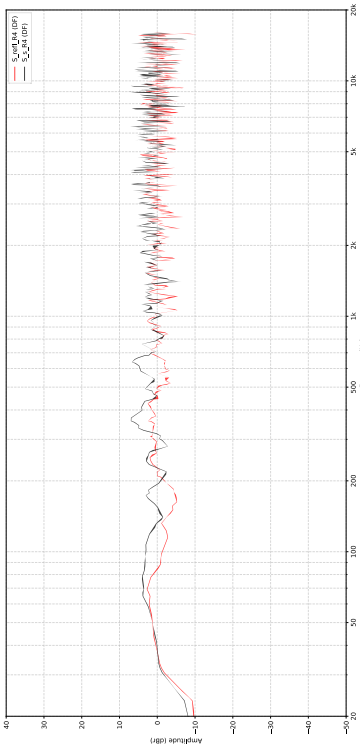


Figure 9.57 S_{refl} and S_s from receiver 4

N.Frequency response plot S_{diff2D} and S_s

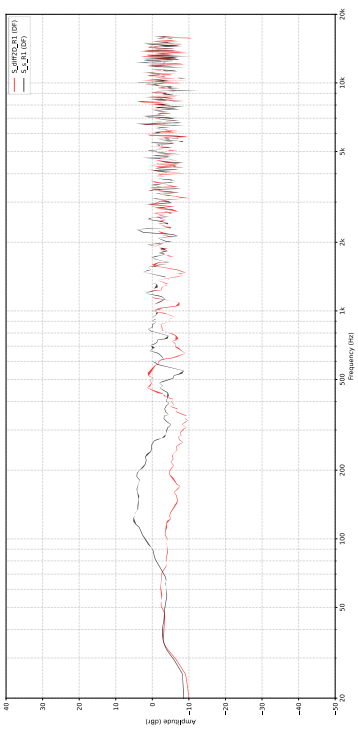


Figure 9.58 S_{diff2D} and S_s from receiver 1

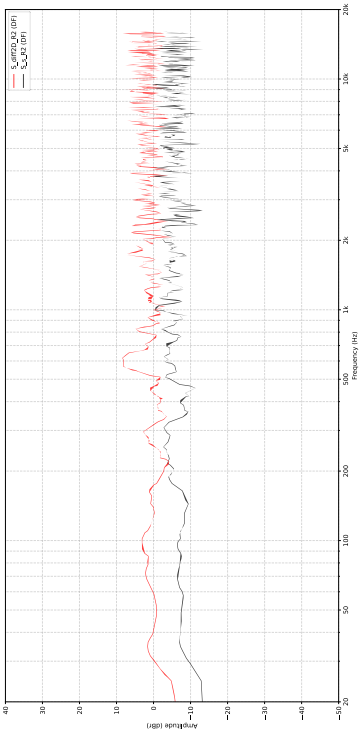


Figure 9.59 S_{diff2D} and S_s from receiver 2

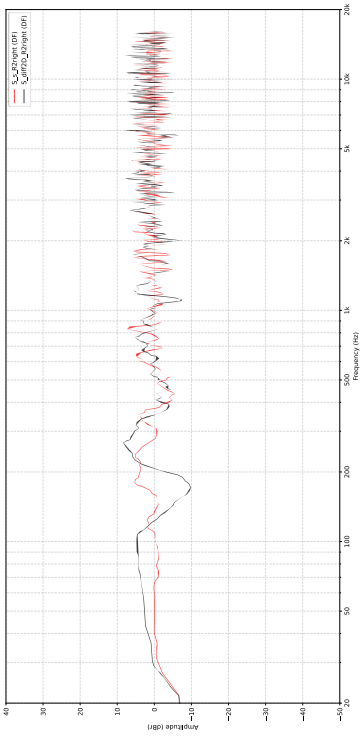


Figure 9.60 S_{diff2D} and S_s from receiver 2 right

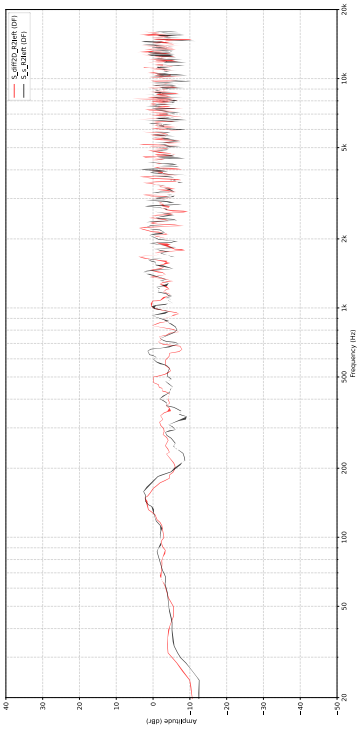


Figure 9.61 S_{diff2D} and S_s from receiver 2 left

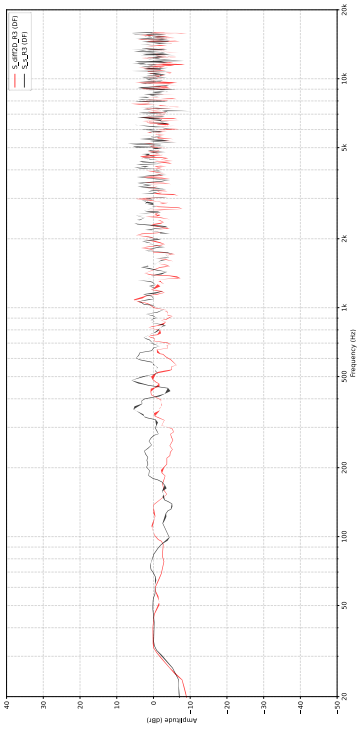


Figure 9.62 S_{diff2D} and S_s from receiver 3

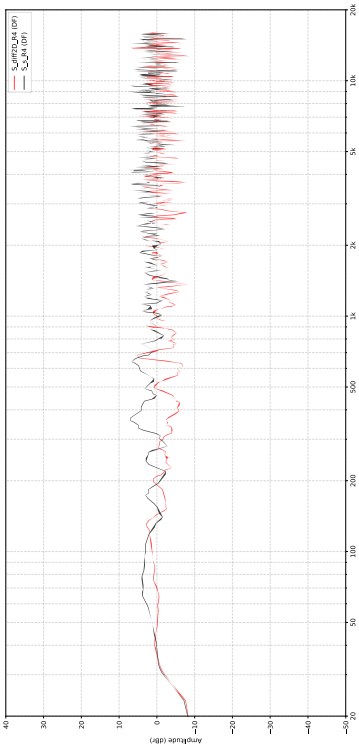


Figure 9.63 S_{diff2D} and S_s from receiver 4

O. Frequency response plot M and S_c

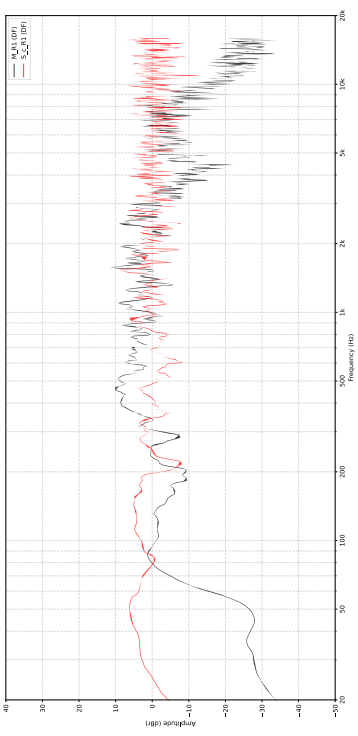


Figure 9.64 M and S_c from receiver 1

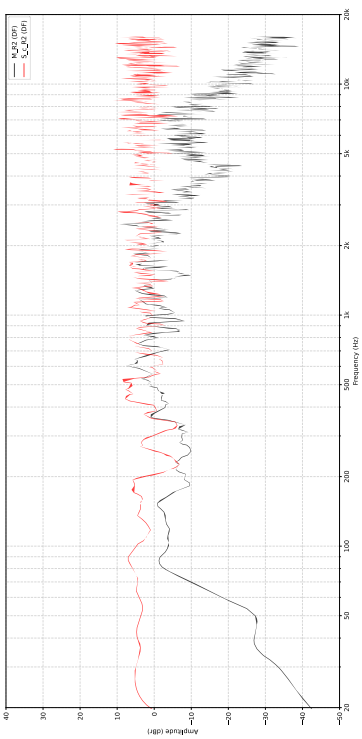


Figure 9.65 M and S_c from receiver 2

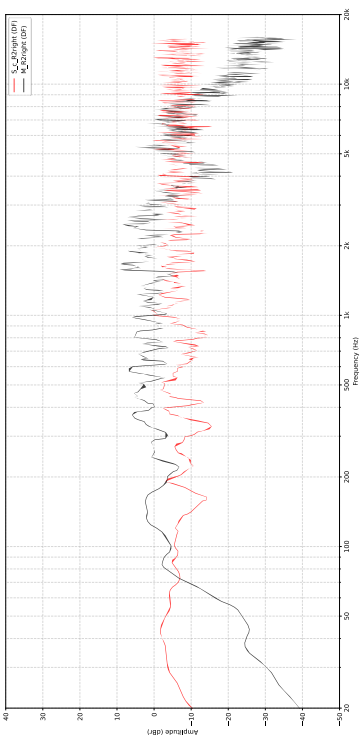


Figure 9.66 M and S_c from receiver 2 right

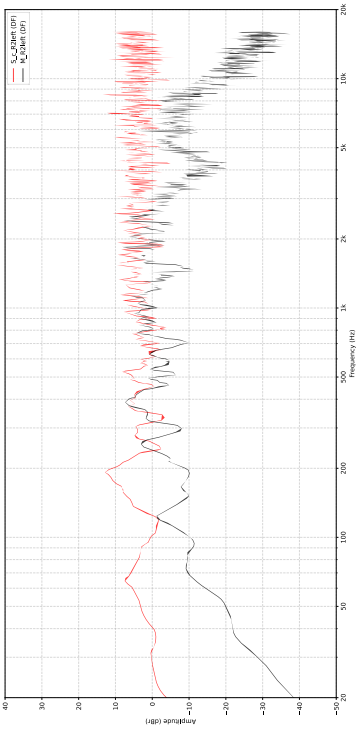


Figure 9.67 M and S_c from receiver 2 left

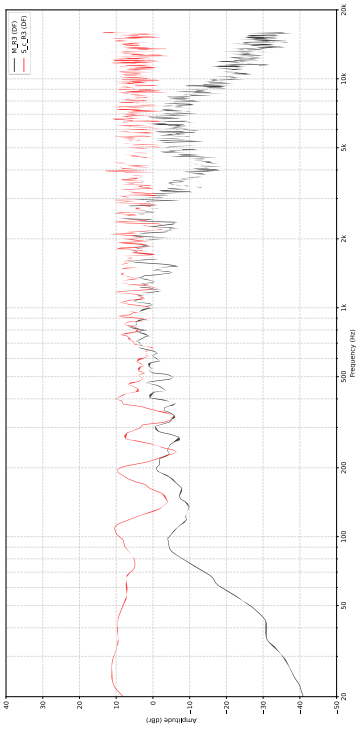


Figure 9.68 M and S_c from receiver 3

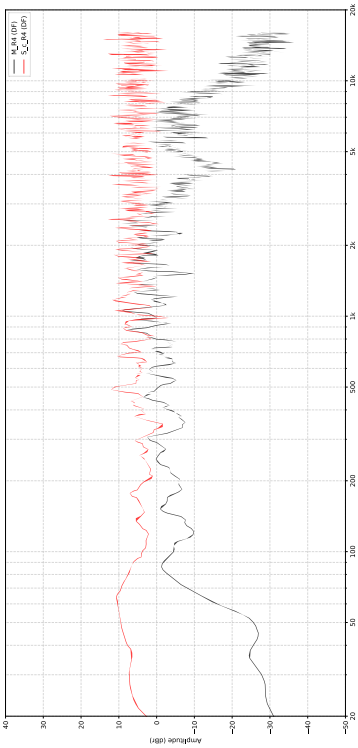


Figure 9.69 M and S_c from receiver 4

P. Frequency response plot M and S_s

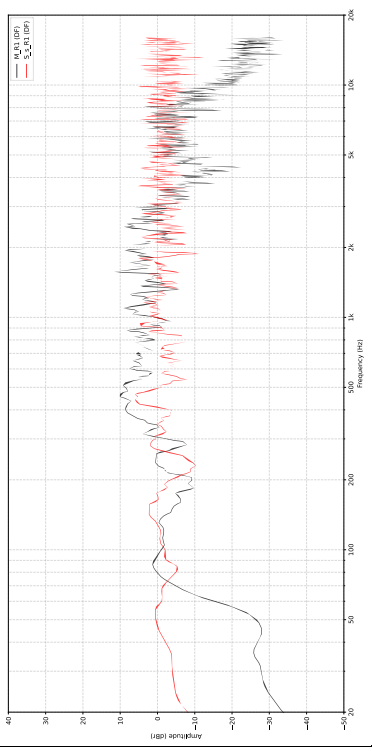


Figure 9.70 M and S_S from receiver 1

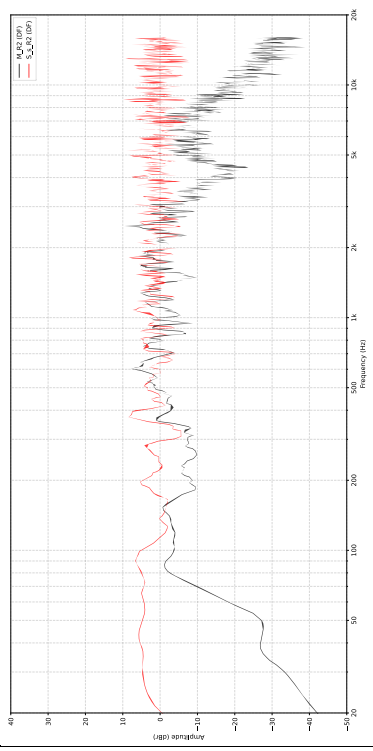


Figure 9.71 M and S_S from receiver 2

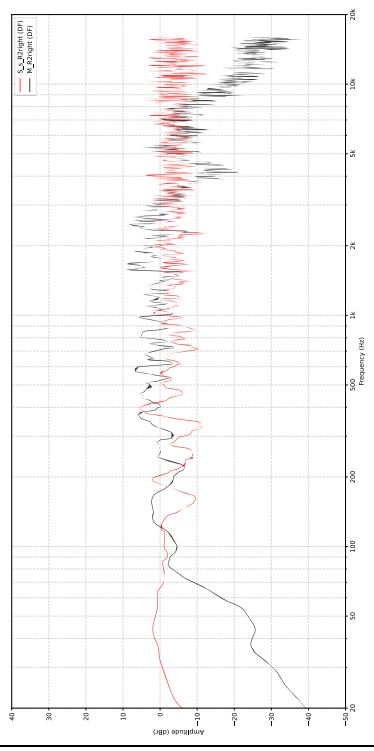


Figure 9.72 M and S_S from receiver 2 right

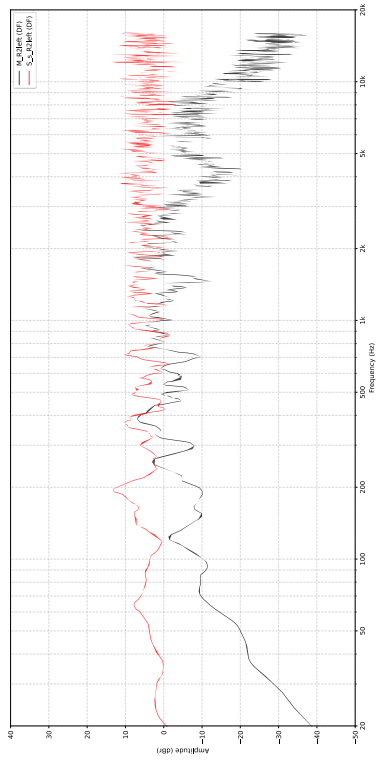


Figure 9.73 M and S_S from receiver 2 left

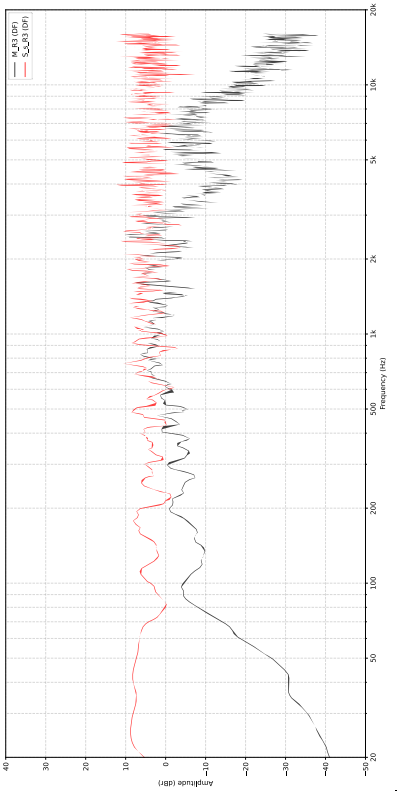


Figure 9.74 M and S_S from receiver 3

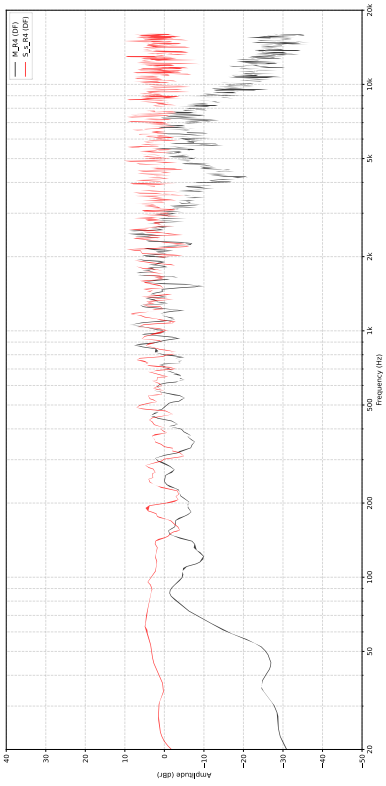


Figure 9.75 M and S_S from receiver 4

Q. Receiver breakdown of stage acoustic (M, S_c, S_s)

A more detailed examination of the discrepancies between measurements and simulations can be found in the graphs in Figure 9.76, which present a comparison of the stage acoustics parameters for each receiver, both under objective measurement conditions and in the simulated results obtained using the complex and simplified models. This analysis, conducted on a per-receiver basis, enables a more profound examination and contextualisation of the average values previously discussed. The initial finding is that there is considerable variability in the values of STearly, STlate, and ST2 among the various receivers. This behaviour is consistent with the literature on stage acoustics, which reports that these parameters were developed precisely to describe the local acoustic conditions perceived by the musician on stage.[26]

The parameters are therefore highly dependent on the receiver's position relative to the surrounding reflective and diffusing surfaces.

Of the three parameters analysed, STlate appears to demonstrate the greatest stability across the different positions. This finding indicates that the contribution of late energy is predominantly influenced by the room's overall acoustic characteristics, while STearly and ST2 demonstrate heightened sensitivity to local geometry and the paths of early reflections. Furthermore, the STlate values generally remain quite negative at all positions considered, indicating that the room effectively limits the accumulation of late energy and prevents the formation of excessive reverberation. In a similar manner, the STearly values remain relatively stable, suggesting the presence of a sufficient contribution from early reflections to provide useful acoustic support to the musicians.

As was evidenced in the analysis of the mean values, a comparison between measurement and simulation demonstrates that the simplified model frequently yields results that are commensurate with those obtained with the complex model, in which the diffusers are explicitly modelled in three dimensions. This finding indicates that, within the framework of geometric-acoustic simulations, the impact of diffusers can be adequately characterised through a statistical depiction of diffusion employing scattering coefficients [54,58]. Finally, it is observed that the discrepancies between the simulated and measured results tend to increase at low frequencies. This behaviour is consistent with the known limitations of simulations based on geometric acoustics: below the Schroeder frequency, the sound field is dominated by modal behaviour and wave phenomena, which are not accurately described by the ray-tracing models used by software such as Odeon[3,52]. Consequently, the agreement between simulation and measurement tends to improve at mid-to-high frequencies, where the sound field exhibits a more diffuse behaviour consistent with the assumptions of geometric-acoustic models.

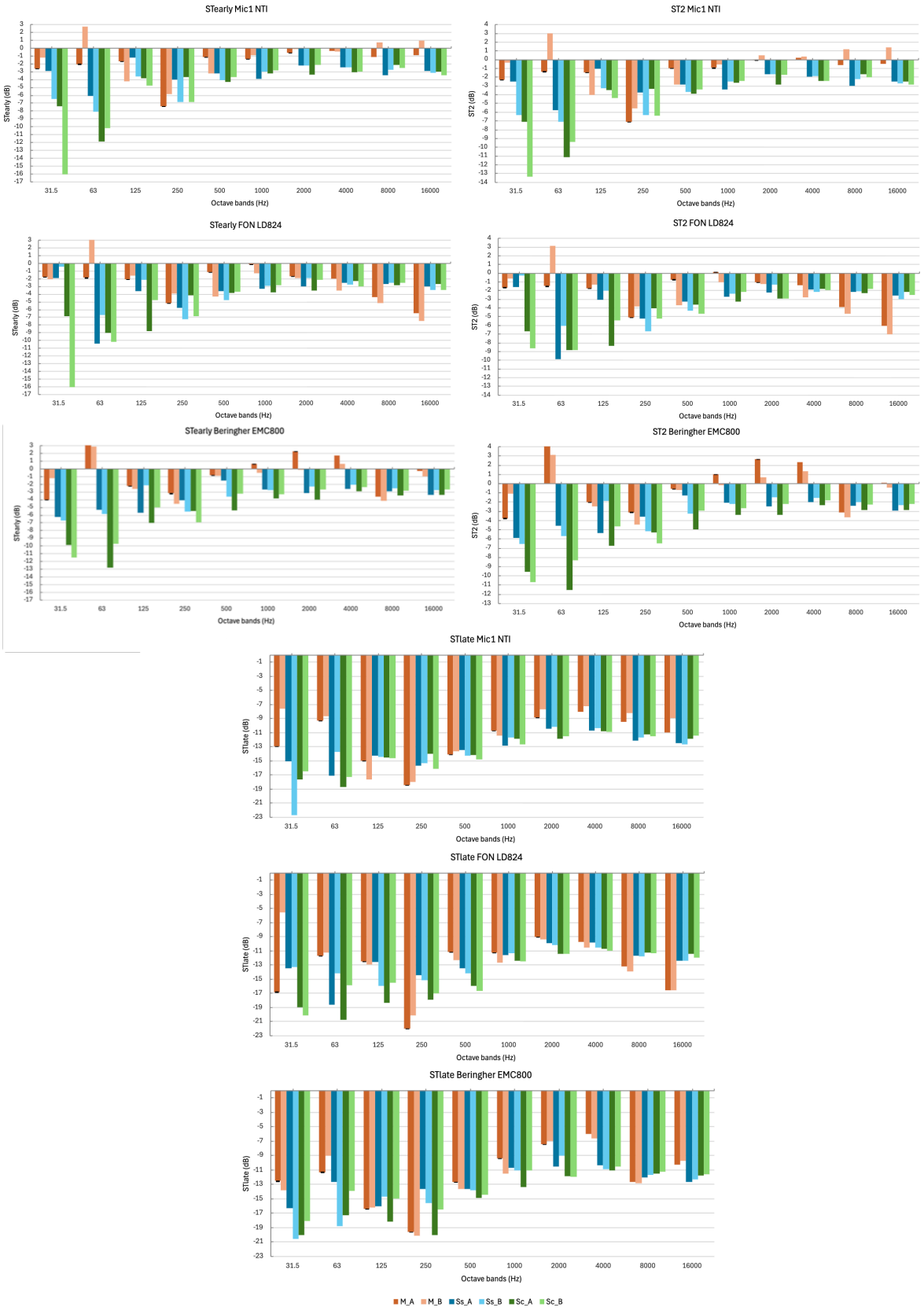


Figure 9.76 Stage acoustic parameters plot for each receiver position of setup A and B

Acknowledgements

Ringrazio la Professoressa Louena Shtrepi, relatrice di questa tesi, che con dedizione mi ha insegnato cosa sia la curiosità scientifica.

Ringrazio il Professore Marco Fringuellino, co-relatore di questa tesi, per avermi trasmesso la passione necessaria per gestire la materia meno tangibile tra tutte: il suono.

Ringrazio Mauro Tabasso, presidente e direttore artistico del Laboratorio del Suono del SERMIG, e il Dipartimento DENERG del Politecnico di Torino, per avermi fornito gli strumenti preziosi e indispensabili per svolgere la ricerca oggetto di questa tesi.

Ringrazio Riccardo Crestani, Andrea Gerbotto, Riccardo Caradonna, Angela Guastamacchia e Lorenzo Lavagna per avermi aiutata con pazienza e competenza.

Ringrazio il Dott. Arch. Filippo Bongiorno, la Dott.ssa Arch. Valentina Delfino e Elia Matino di Open Architecture Studio che mi hanno accolta nel loro studio, mettendo a disposizione le loro risorse e facendomi sentire quasi una BIM designer.

Ringrazio Gabriele Monteleone, perché il suo affettuoso aiuto e la sua dolcezza mi ha permesso di superare ostacoli che sembravano insormontabili.

Ringrazio tutti i colleghi incontrati in questo percorso: quelli con cui ho preparato esami e svolto progetti di gruppo, ma anche quelli con cui ho scambiato solo un saluto nei corridoi.

Ringrazio le otto bocciature di Analisi 1 e le sei bocciature di Fisica 1, perché hanno confermato il dubbio di tutti: dopo l'ansia, il modo per farcela lo trovo sempre.

Ringrazio Gemma, perché è l'amica che il primo giorno di università non sopportavo, ma l'unica che c'è anche oggi a festeggiare questo grande traguardo insieme.

Ringrazio la musica, che in questi anni ho allontanato e poi ritrovato, perché rimane l'amante da cui torno sempre e a cui non saprei mai dire di no.

Ringrazio la mia famiglia, sempre un porto sicuro.