Validation of room acoustic simulation models

Von der Fakultät für Elektrotechnik und Informationstechnik der
Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades eines

DOKTORS DER INGENIEURWISSENSCHAFTEN

genehmigte Dissertation

vorgelegt von

Dipl.-Ing.
Lukas Aspöck

aus Duisburg, Deutschland

Berichter:

Universitätsprofessor Dr. rer. nat. Michael Vorländer
Universitätsprofessor Dr. Stefan Weinzierl

Tag der mündlichen Prüfung: 10. Juli 2020

Lukas Aspöck

Validation of room acoustic simulation models

Logos Verlag Berlin GmbH

λογος

Aachener Beiträge zur Akustik

Editors:
Prof. Dr. rer. nat. Michael Vorländer
Prof. Dr.-Ing. Janina Fels
Institute of Technical Acoustics
RWTH Aachen University
52056 Aachen
www.akustik.rwth-aachen.de

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de .

D 82 (Diss. RWTH Aachen University, 2020)

ISBN 978-3-8325-5204-6

ISSN 2512-6008

Vol. 35

Logos Verlag Berlin GmbH
Georg-Knorr-Str. 4, Geb. 10,
D-12681 Berlin

Tel.: +49 (0)30 / 42 85 10 90
Fax: +49 (0)30 / 42 85 10 92
http://www.logos-verlag.de

Abstract

Already in ancient times, the goal existed to design appropriate environments to transport acoustic signals from a sound source to an audience. Achieving this goal requires an understanding of room acoustic effects and, in the best case, the ability to accurately predict them. Simulations based on the concepts of geometrical acoustics have been developed and extensively investigated for more than 50 years. However, even today, there is still insufficient knowledge about the uncertainty and validity of the acoustical simulation results.

In order to offer developers and researchers a possibility to comprehensively validate the results of their room acoustic simulation, a database with eleven acoustic scenes was developed, documented and published. The scenes cover both simple environments, where individual acoustic effects such as reflection or diffraction are isolated, and complex room situations such as a concert hall. In addition to a detailed description and the definition of input data, acoustic measurements were carried out for all scenes, which serve as a reference for simulations.

Based on this data, a first study is presented which was designed following the concept of three previously conducted round robin experiments on room acoustic simulation. Here, the simulation results of up to six participants, who were not informed about measurement results in advance, are compared with each other and with the results of the corresponding measurements. The evaluation of the simple scenes revealed various weaknesses of the simulation software. In case of the complex scenes, acceptable results were observed in the mid-frequency range. In a second study it was determined how much room acoustic simulations deviate from measurements if the user knows the measured results beforehand and input data of the simulation is adjusted manually or systematically.

This work demonstrates which deviations from actual, measured values can be expected when simulating either unknown or known rooms. It furthermore provides a basis for the revision of room acoustic simulation software and for the improvement of the determination and selection of correct input data.

Kurzfassung

Bereits in der Antike existierte das Ziel, geeignete Umgebungen zu gestalten, um akustische Signale von einer Quelle zu einem Publikum zu transportieren. Dazu müssen raumakustische Effekte verstanden werden und, im besten Fall, diese auch genau hervorgesagt werden. Simulationen, die auf den Konzepten der geometrischen Akustik basieren, werden seit mehr als 50 Jahren entwickelt und ausgiebig untersucht. Dennoch ist auch heutzutage über die Unsicherheit und die Validität der Resultate der Simulationen nicht ausreichend bekannt.

Um Entwicklern und Forschern eine Möglichkeit zur Validierung ihrer Simulationen zu bieten, wurde eine Datenbank mit elf akustischen Szenen entwickelt, dokumentiert und veröffentlicht. Die Szenen decken sowohl einfache Umgebungen, in denen einzelne akustische Effekte wie Reflexion oder Beugung isoliert werden, als auch komplexe Räume wie einen Konzertsaal ab. Neben einer detaillierten Beschreibung und der Definition von Eingabedaten wurden für alle Szenen akustische Messungen durchgeführt, die als Referenz für die Simulationen dienen.

Basierend auf diesen Daten wird eine erste Studie vorgestellt, die nach dem Vorbild der drei bisher durchgeführten Round Robins der raumakustischen Simulation konzipiert wurde. Dabei werden die Simulationsergebnisse von bis zu sechs Teilnehmern, denen im Vorfeld die Messergebnisse nicht bekannt waren, miteinander und mit den Messergebnissen verglichen. Die Auswertung der einfachen Szenen offenbarte Schwächen der Simulationssoftware, zeigte aber bei den komplexen Szenen akzeptable Ergebnisse im mittleren Frequenzbereich. In einer zweiten Studie wurde ermittelt, wie stark raumakustische Simulationen von Messungen abweichen, wenn die Messergebnisse im Vorfeld bereits bekannt sind und die Eingabedaten der Simulation manuell oder systematisch angepasst werden.

Diese Arbeit demonstriert, welche Abweichungen von gemessenen Werten zu erwarten sind, wenn unbekannte oder bekannte Räume simuliert werden. Zudem dient sie als Grundlage für die Überarbeitung von Simulationssoftware und für eine verbesserte Bestimmung und Auswahl von korrekten Eingangsdaten.

Contents

Acronyms

BEM	boundary element method
BRAS	Benchmark for Room Acoustical Simulation
BRIR	binaural room impulse response
C80	clarity (room acoustic parameter)
CR	complex room
DT	decay times
EDT	early decay time
EDC	energy decay curve
FABIAN	Fast and Automatic Binaural Impulse response AcquisitioN
FEM	finite element method
GA	Geometrical Acoustics
IS	image source
IR	impulse response
JND	just-noticeable difference
LDT	late decay time
HATO	head-above-torso orientation
HRIR	head-related impulse response
HRTF	head-related transfer function
LTI	linear time-invariant
MDF	medium-density fibreboard
RAVEN	Room Acoustics for Virtual ENvironments
RIR	room impulse response
RS	reference scene
RTF	room transfer function
RT	ray tracing
SAQI	Spatial Audio Quality Inventory
SD	standard deviation
T20	reverberation time (room acoustic parameter)
T30	reverberation time (room acoustic parameter)

1

Introduction

Sound is an important part of every person's daily life. Due to the complex composition of sound signals influenced by numerous physical factors, sound can not only vary greatly in level and frequency, but also in its effect on us human beings. Sound can bring joy, but it can also disturb us at work and in the worst case even severely affect our health. The environment in which we are perceiving sound sources is highly relevant. While, from an evolutionary point of view, humans have spent most of their active time outdoors, nowadays it is often the case that humans spend most of their time indoors. Therefore there is a need to adequately design the acoustics of indoor environments, e.g., by applying models to predict these room acoustics before a building is constructed.

With the immense progress in technology, various types of simulation models have emerged and are nowadays applied in software for the assessment of rooms ranging from car interiors and classrooms to large auditoriums and concert halls. Using concepts from signal processing, a room can be characterized by an impulse response, which sufficiently describes the acoustic properties of a room from a technical point of view. If however, the realistic experience of being in a room should be simulated, one also has to account for our multimodal perception and with respect to our hearing, especially for our binaural sensation.

Despite enormous research efforts in the field of room simulation using the concept of Geometrical Acoustics (GA), it is still a challenge to create authentic simulations which, if assessed in psychoacoustic studies, can not be distinguished from the corresponding real situations. In order to investigate the validity of room simulations based on geometrical acoustics, this work presents a database intended for the validation of simulation algorithms. Furthermore two studies are presented comparing simulated and measured data in a technical analysis.

The remaining part of this chapter includes an overview of related work including publications from the past 50 years. In Chapter 2, a brief introduction to room

acoustics is given including methods to measure and simulate the acoustical characteristics of a room. Chapter 3 presents and discusses the scene database and Chapter 4 deals with a round robin comparison including different simulation algorithms. In the study presented in Chapter 5 simulations created by the author of this work are compared to the corresponding measurements. Chapter 6 summarizes this work and gives an outlook with respect to future research.

Overview of related work

Acoustical simulations intended for the application in room acoustics have been extensively researched in the past decades. With an increasing number of applications areas, ranging from architecture and planning [86, 184, 71] to games and virtual reality [177, 132, 149], the need for robust, accurate simulations as well as plausible or authentic auralizations today is greater than ever. Plausibility or authenticity [96, 39] of auralizations including room simulations is also particularly crucial if applied in psychoacoustic experiments [27, 190, 61, 108] and for hearing aid research [60, 120].

With respect to different models, their implementation and extensions towards wave propagation effects such as diffraction, the reader is referred to the paper recently published by Savioja and Svensson [148] which includes a very extensive overview of GA based simulation models. A comprehensive overview and a comparison between different wave-based simulation models is published by Sakuma et al. [144].

This section will primarily cover validation studies of GA models and software, with a focus on the previous round robin investigations. Additionally some more recently proposed simulation models and related studies are presented.

Early validation studies

While publications of the first implemented computer models [85, 2, 28, 174] only presented and discussed implementations, at the end of the 1980s an increasing number of simulation models were validated in case studies comparing simulated data to measurements conducted either in real rooms or in scale models.
In 1989, Ondet and Barbry validated a ray tracing approach for noise level calculations in a workshop including many scattering objects and concluded that the comparison with the measurement "*is extremely satisfactory*" [117]. In this study, however, no other room acoustic parameters were evaluated or discussed. Hodgson later found "*excellent agreement*", also for energy decay curves, in a

study using the same ray tracing model [69]. In another publication [70], Hodgson compared several prediction models for the level calculation in fitted rooms, mostly image source based but also including the ray tracing model of Ondet and Barbry. This study concluded that this ray tracing model "*far outperforms the others tested and appears to be the best choice for fitted-room prediction*".

Kuttruff presented one of the first computer-based implementation of auralizations using simulated binaural room impulse responses [86]. In this publication it is claimed that "*subjective comparison of the music samples showed almost perfect agreement, samples processed with the auralization filter and the natural impulse response are nearly undistinguishable*" without providing details of this listening experiment.

Naylor and Rindel compared results of an early version of the ODEON simulation tool (2.0) with real room and scale model measurements and reported that "*quite good agreement between measured and calculated results can be achieved without enormous computing effort*" [110]. In this study, values for measured and simulated room parameters for different frequencies are presented and discussed. In case of the real room, the deviations of the room parameters are below or just slightly above the corresponding just-noticeable difference (JND) and are regarded in general as *satisfactory*. Higher deviations are observed for the comparison with scale model measurements, which are mainly attributed to the uncertainty of the measurement and the challenge to define adequate absorption coefficients for the simulation. In another study of the early 1990s which compared scale model measurements with an image source model, Kleiner et al. [83] reported "*reasonable agreement with regard to measurement uncertainty*" and also highlighted the importance of input data: "*the choice of materials and measurement technique for the physical scale model as well as input data for the computer model are critical in this respect*". Additionally, the *validation problem* is explained by stating that "*absolute accuracy is at present some way off*" and that validation is achieved, when the comparison of impulse responses and parameters to a given precision are satisfactory for a large number of source and receiver position combinations.

Another study comparing different GA-based approaches to measured values was presented by Lam [90]. Here, three different methods for diffuse reflection calculation were investigated and the question was raised, in how far the same *diffuse-reflection* coefficients can be used for different models. This was confirmed, at least for typical performance spaces with small aspect ratios. For the results of a real multipurpose auditorium, acceptable deviations were found for level

("*seemed appropriate*") and reverberation time (" *... predictions of the average reverberation time were reasonably good*"). For more sensitive parameters such as clarity and the early lateral energy fraction, substantial deviations and differences between the scattering models were discovered.

Round robin investigations

In comparison to previous studies conducted by Hodgson [70] and Lam [90], which also evaluated and multiple simulation models, in round robin comparisons, different users of the prediction models have to apply the provided information and input data in their preferred simulation software. Thus, the user of a room simulation software is also part of the comparison. This includes individual design decisions and potentially also mistakes with respect to the preparation of the scene's input data and the configuration of the simulation. Regarding this, Bork stated that in round robins [29], "*not only the quality of simulation software influences the results but also the skill of the user as an acoustician*".

With more simulation models and software emerging in the early 1990s, the first round robin on room acoustic simulation was initiated in 1993 by Vorländer [176]. The most essential facts of this investigations are listed in Table 1.1. No 3D model of the room was included in the scene description, instead drawings along with material descriptions were provided. This lead to large discrepancies among the participants and deviations to the measured values in the first phase. In phase II, the impact of the user was reduced as much as possible by providing input data. Discrepancies among the participants decreased, but were still substantial in comparison to the standard deviation of the measured values, which were conducted by seven groups. In general, three programs were identified to be "*unquestionably reliable in the prediction of room acoustical parameters*". A reason for deviations of the applied software was attributed to the neglection of attenuation at grazing incidence over seat rows. Details of the second round robin on room simulation software [29] are listed in Table 1.2. Here, the same concept using two phases was applied, but for a considerable larger test room, the ELMIA concert hall in Sweden. The evaluation was also extended to six octave bands and nine different room parameters. The first phase was considered for the participants to be a "*first guess*", which is heavily influenced by the experience and skill of the participants with respect to acoustic design of rooms. In the second phase, which was suppose to highlight the real differences of the software, still seven out of 13 submissions were based on 3D models which were different from the provided model. As no in situ measurements in the hall could be conducted, datasets for absorption and diffusion coefficients were adopted from literature

Table 1.1: Details of the first round robin on room acoustic simulation (1993-1994), results published in 1995 [176]

Room	PTB lecture hall, V=1,800 m^3, 274 seats
Positions	Two sound source and five receiver positions
Geometry	Drawings of the room: ground plan and side view
Phase I	Individual absorption data
Phase II	General absorption data (provided)
Submissions	16
Algorithms	Image sources (2), ray tracing (7) or hybrid models (7)
Evaluation	Eight room parameters for 1 kHz octave band

and based on estimations, respectively. During phase II, participants were aware of the measured data, but Bork stated that "*it is assumed that none of them did change his software algorithms in order to get closer to the measurements*" [29].

The result analysis revealed highest deviations for parameters in the 125 Hz frequency band, which were mainly attributed to inaccurate absorption coefficients for the seating. The mostly consistent input data of phase II also lead to a substantial reduction of the variations among the submissions, but due to erroneous absorption coefficients, reverberation times and early decay times averaged across evaluated positions and frequencies consistently deviated around two JNDs from the measured values. A systematic overestimation of clarity and definition values for 125 Hz was also discovered, but could not be explained in this study.

Table 1.2: Details of the second round robin on room acoustic simulation (1996-1998), results published in 2000 [29]

Room	ELMIA hall, V=11,000 m^3, 1100 seats
Positions	Two sound source and six receiver positions
Geometry	Photos, drawings and 3D model
Phase I	Description of surface materials
Phase II	Absorption and diffusivity data (estimated)
Submissions	16 (phase I), 13 (phase II)
Algorithms	Mostly combined image source/ray tracing algorithms
Evaluation	Nine parameters for six octave bands (125 Hz - 4 kHz)

Results of the third round robin on room acoustic simulation, also organized by Bork, were published in 2005 [30, 31]. In this investigation, a recording studio, including diffusing elements on one wall and the ceiling, was selected as the test room. The volume of the room is 400 m^3, more details of this study are listed in Table 1.3. In contrast to the previous investigations, all input data was provided

Table 1.3: Details of the third round robin on room acoustic simulation (1999-2002), results published in 2005 [30, 31]

Room	PTB recording studio, V=400 m^3
Positions	Two sound source and six receiver positions
Geometry	Different 3D model for each phase
Phase I	Simple room model with seven walls, homogeneous absorption & scattering data for all walls & frequencies
Phase II	Detailed model, diffusing elements as planes. Frequency dependent absorption and scattering coefficients.
Phase III	Detailed model, including modeling of diffusing elements.
Configurations	Room curtains were either closed or open (phase II & III).
Submissions	21 (9 programs), 6 selected programs for phase II & III
Algorithms	Mostly combined image source/ray tracing algorithms
Evaluation	Nine parameters for six octave bands (125 Hz - 4 kHz)

to the participants (3D models, absorption and scattering coefficients). No in situ measurements were conducted, instead tabulated and estimated data was provided with the exception of the diffusing elements, for which scattering and absorption coefficients were determined in reverberation chamber measurements.

Phase I was designed to be a simple reference case in order to compare the applied software. Here 14 out of 21 participants delivered reasonable data resembling the results obtained by Sabine's equation. The results of the other participants substantially deviated from these values or failed to correctly account for air absorption. For phase II and III, only the results of six commercial simulations were evaluated. In comparison to the previous round robins, these results showed less variation among the participants and at least for position-variable parameters such as clarity, good agreement with the measured values. Simulations also accounted for varying configurations of curtains in the room. For the reverberation times, however, deviations were consistently above 2 JNDs, even for the 1 kHz frequency band. In general, more deviations were observed for the lowest frequency bands. These deviations were attributed to missing diffraction simulation and to the inability to process complex wall impedances in the applied simulations.

A round robin comparison for wave-based methods, not only with respect to accuracy, but also to efficiency of the simulation, was initiated by Sakuma et al. [145] in 2002. The related publication, however, only presented the framework and some trial calculations. Later a benchmarking platform for numerical methods[1] was presented [143, 119]. It contains several scenarios related to environmental and architectural acoustics, some of them are also discussed in [144].

Another round robin related to room acoustics was conducted by Katz [81], not examining simulations results, but different implementations for the room acoustic parameter evaluation based on measured RIRs. This study revealed a substantial degree of variation between the analysis methods, especially for the lowest evaluated frequency band (125 Hz), but also with variations close to the subjective difference limen in the 1 kHz frequency band. A follow-up study by Cabrera et al. [43] with a focus on the reverberation time showed that variations of nine implementations have decreased in comparison to the previous study.
Recently a round robin on coupled rooms was conducted by Weber and Katz [181]. This comparison included a simple coupled room situation with homogeneous boundary conditions, for which 11 participants (using three wave-based simulations and seven different GA-based simulation tools) provided impulse responses for the individual rooms as well as for the coupled room situation. The results demonstrated that the tested tools are able to simulate coupled room effects, and are, to some extent, in agreement with the theoretical model [47].

Room acoustic simulation models: the last 20 years

The review of publications with respect to room acoustic simulation of the last 20 years reveals a shift from parameter based evaluation towards perceptual validation of simulations and their corresponding auralizations. Lokki and Järveläinen, for example, presented a perceptual evaluation of real-time auralizations of static and dynamic situations based on the DIVA auralization system [147]. Choi and Fricke [44] also compared simulations with measurements of two concert halls based on parameters and listening experiments. Pelzer et al. [122] conducted a study comparing measured and simulated results based on a wave-based model combined with a GA model and also included a technical analysis of binaural room impulse responses, "*which indicated a very promising agreement*", but was not evaluated in listening experiments.

[1] http://news-sv.aij.or.jp/kankyo/s26/AIJ-BPCA/A0-1F/index.html (Accessed: March 2020)

Other studies have a specific focus, such as the general uncertainty of room simulations [178], comparisons of parameters evaluated for a full audience area [146], a detailed analysis of impulse responses and reflections in addition to traditional parameter analysis [170] or the inclusion of source and receiver directivity data in boundary element method simulations [65].

Despite the trend towards auralizations and perceptual investigations, improvements and extensions of basic GA models still have been the subject of research, e.g., the implementation of radiosity methods [114] or inclusion of phase effects and complex surface impedances [5, 101, 32]. A model to transform absorption coefficients into complex surface impedances in order to use them in software applying extended GA models, was recently proposed by Mondet et al. [105]. An interesting observation of the more recent studies is that not all simulations are compared with measurements. In the study about acoustical radiosity by Nosal et al. [114], the predicted results are compared to analytical solutions of a spherical enclosure. Aretz et al. [5] compared the presented extended image source model for small rectangular rooms with corresponding finite element simulations. Such validations, however, can only be conducted for very simple room geometries or in case of the wave-based simulations, are limited to lower frequencies. In a promising approach presented by Wang et al. [179], wave-based simulations using the discontinuous Galerkin method were referenced to both, analytical solutions and a real room measurement. This simulation was also investigated for a recently designed benchmark room [58].

Another direction, which is typically considered when real-time simulations of sound fields are desired, is the simplification of GA models, e.g. by using shoebox models instead of detailed geometries in order to create plausible and efficient room auralizations [185, 186]. Also, the application of machine learning and neural networks for efficient room acoustic rendering has been proposed [53, 166, 167].

In general, it is unfortunate that for most mentioned studies the applied scene data, corresponding detailed documentation and reference measurements are not publicly available. More comprehensive evaluations and meaningful comparisons of simulations would be possible if more well documented data is available. In addition to the above mentioned database initiated by Otsuru et al. [119], also the European Acoustics Association hosts a platform[2] intended for validation of wave-based simulations. Mathematical scene descriptions are provided, however, no measured or analytical references are included.

[2] https://eaa-bench.mec.tuwien.ac.at/main/ (Accessed: March 2020)

When it comes to the definition, selection and measurement of input data, especially for complex real world scenarios, the challenge of describing adequate input data for room simulation remains, as discussed with respect to geometry in [8] and with respect to the boundary conditions in [178, 80]. Therefore, little has changed in the evaluation of room acoustic simulations compared to the studies conducted 30 years ago. Results can still be considered acceptable, but despite intensive research efforts and numerous new models, the case studies only show minor improvement in the case of complex real-world room scenarios.

2

Room acoustics: Measurement and simulation

This chapter briefly introduces the underlying concepts of room acoustics, room acoustical simulation and its validation. For more extensive explanations of theories and models, the reader is referred to the corresponding textbooks with respect to sound propagation [87, 141], room acoustics [47, 18, 141, 88], boundary conditions [46], audio signal processing [183] and room simulation and auralization [177]. While for room acoustic simulation no standardized methods have been established (yet), the measurement of room acoustic parameters is standardized in ISO 3382 [76]. For room simulation software, the methods applied for acoustical measurements and its evaluation procedures are adopted. The following sections cover the basics of sound propagation in rooms, room acoustic measurements and prediction models.

2.1 Sound propagation in rooms

Sound is typically described by acoustic waves traveling through a medium. In this work, only sound propagation for frequencies between 20 Hz and 20 kHz through air at a speed of roughly 343 m/s, emitted by electroacoustic transducers, is considered. In rooms, such sound waves experience various physical effects such as interference, diffraction, reflection, scattering or air attenuation. How sound waves arrive at a receiving position in a room can be described by the room impulse response (RIR). This concept is adopted from system theory, and considers a room as a linear time-invariant system (LTI-system). Thus, a RIR describes the path from the sound source to the receiver in a room, with the option to either include or exclude characteristics of both, the sound source and the receiver. For an evaluation in the frequency domain, a room transfer function (RTF) can be calculated by applying a Fourier transform [118].
In the domain of digital audio processing, RIRs are typically sampled at a rate of at least twice the maximum audible frequency of humans (20 kHz), according to the Nyquist–Shannon sampling theorem [116]. In this work, a sampling rate of 44.1 kHz is considered.

Measurement and evaluation of room acoustics

The concept of using the RIR to describe the acoustics of a room is particularly useful if reliable and valid measurement techniques exist to determine the impulse response. An easy way to obtain a RIR, which is also included in the measurement standard ISO 3382 [76], can be achieved by introducing an impulsive sound, e.g., a balloon pop or a gun shot, into a room and measuring the result with an audio recording device. This method lacks reliability and also validity, especially if the bandwidth of the measurement should correspond to the audible frequency range [79, 102]. More sophisticated methods to determine the RIR are omnidirectional measurement speakers, which emit maximum-length sequences (MLS) [136] or exponential sine sweeps [54, 103]. For conducting room acoustic measurements, the latter measurement signal is considered as the preferred option as it is easier to achieve higher levels for the signal-to-noise ratio [55, 62]. The concept of a RIR measurement using an omnidirectional measurement speaker and an omnidirectional measurement microphone is visualized in Fig. 2.1.

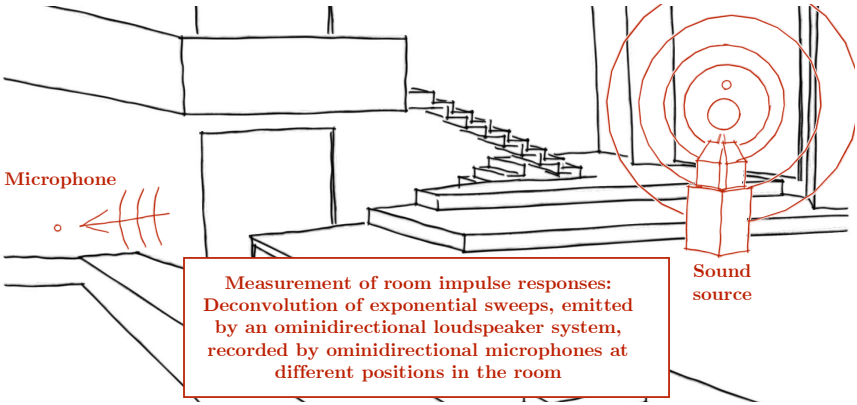


Figure 2.1: Schematic presentation of a room acoustic measurement

To calculate room acoustic parameters of a RIR, the *Schroeder* curve is processed by backward integration of the squared RIR [154, 76]. This is typically done either for one-third octave or for full octave frequency bands. The most popular room acoustic parameter is the reverberation time, which was first described by Sabine [142] who is also referred to as the initiator of room acoustics as a field of research. The reverberation time describes the duration it takes until the level of a sound in the room is decreased by 60 dB, typically calculated by linear regression for an interval of the *Schroeder* curve.

Many other quantities exist, many among them are purely based on the squared RIR such as sound strength, clarity or the bass ratio. Other parameters such as the lateral energy fraction or the interaural cross-correlation coefficient require additional measurements with a figure-of-eight microphone or with a dummy head, respectively [76]. For the sake or brevity, the parameter evaluation in this work is mostly restricted to the reverberation time (T20 or T30), the early decay time (EDT) and the clarity parameter (C80).

2.2 Prediction models

Several options exist to analyze or predict the acoustical behavior of a room, ranging from simple equations for the prediction of parameters to scale models and simple-to-complex simulation models in order to generate full length RIRs. The most simple and established reverberation time equations of *Sabine* and *Eyring* only require the volume and the equivalent absorption area of the room. The latter is typically calculated using tabulated values for the different surface materials of the investigated room. Interestingly, these rather simple equations, which assume ideal diffuse field conditions, are still applied in practice today. Many other models have been proposed which, in some cases, turned out to be a more accurate predictor in challenging conditions such as for inhomogeneous absorption distribution in the room [56, 3, 111]. For other room acoustic parameters, regression models based on empirical data are listed by Gade in [141]. Here parameters such as strength or clarity are based on geometric input and on more simplified estimations derived from the diffuse field theory.

An option to evaluate and predict the acoustics a room in the planning phase are scale model measurements, which were first presented by Spandöck in 1932 [163]. While in the early days of this method, the airborne acoustic wave propagation in these models was even replaced by water surface waves or by light, which came with various limitations and restrictions with respect to the evaluation, conducting room acoustic measurements with small eletroacoustic transducers and microphones in the scale models provides sufficiently accurate results in order to predict the sound propagation in the real room [88]. Even today, especially for expensive concert hall projects, scale models with scaling ratios of 1:10, 1:20 or 1:50 are still being used, allowing an evaluation of a RIR before a concert hall is built.

In theory, the most accurate way to predict the sound field in a room is the application of numerical, wave-based simulation models [144], such as the finite-difference method, the finite element method (FEM) or the boundary element

method (BEM) [165, 64]. For the application in room acoustics, these models find numerical solutions of the three-dimensional wave equation, a partial differential equation. Such models, however, require detailed input data with respect to the boundary conditions, typically described by complex valued surface impedances, and the room mesh. The applied mesh discretization determines the upper frequency limit of the simulation and strongly impacts the required calculation time, which is normally in the range of multiple hours or even multiple days, especially if a simulation up to highest audible frequencies is desired. For this reason, and despite increasing available computational power of modern computers and cluster systems, most room simulations are nonetheless conducted with software based on GA models.

2.2.1 Room simulation based on geometrical acoustics

In the field of GA, sound waves are modeled as rays. While first computer simulations were not described before the paper of Krokstad et al. [85] in 1968, the concept of ray acoustics is the basis for describing the mean free path length and thus also for development of the reverberation time equation. Replacing sound waves by rays can represent a severe simplification of the physical behavior, but eventually leads to very efficient numerical simulations accounting for room geometry and the specific distribution of different surface materials in the rooms. These models are therefore more flexible than the simple equation based prediction models, which only provide reliable results for ideal, uniform conditions [155]. In addition to the computational advantage over wave-based approaches, the input data of GA simulations is easier to prepare with respect to the room geometry and the boundary conditions, making GA software more popular, especially for practitioners.

In GA software, sound sources are modeled as point sources, emitting energy which arrives at the receiving point either on a direct or reflected sound path. The basic concept of a GA-based simulations is sketched in Fig. 2.2. Sound rays starting at the sound source are either reflected specularly or scattered at the room's surfaces, represented by a 3D model. Two sound path examples are visualized in the sketch: The direct sound from source to receiver, and a first order specular reflection on the floor of the stage. The reflection type and direction are determined based on scattering or diffusion coefficients, the energy loss of a reflected ray is based on the absorption coefficients of the materials which are assigned to each surface of the 3D model.

Different approaches exists to determine these sound paths, ranging from classical ray tracing (RT) [85] and the image source (IS) model [2] to variants of ray tracing techniques such as beam-tracing [57] and combinations of both in hybrid

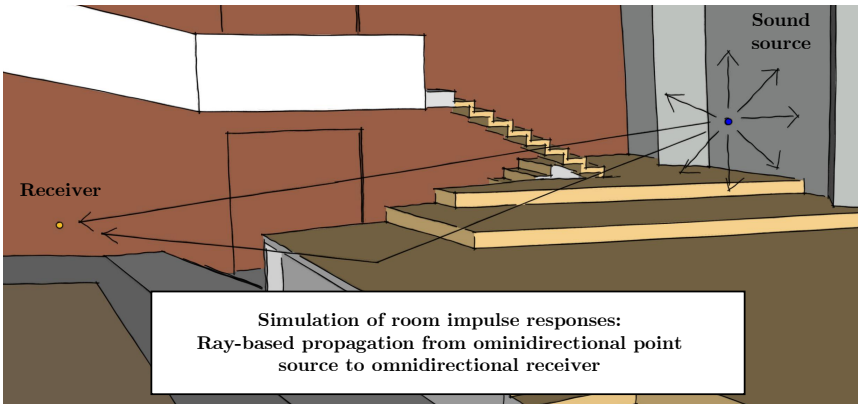


Figure 2.2: Schematic presentation of a room acoustic simulation

simulation models [174]. The hybrid approach is also applied in most commercially available tools for room acoustic simulation, such as CATT-Acoustic [49], ODEON [115] or EASE [1].
While the image source model directly contributes reflections to the impulse response, RT implementations are typically used to generate an energy histogram (also called echogram), which subsequently need to be processed in a filter synthesis process [89]. The application of head-related transfer functions (HRTFs) in this filter synthesis [151, 7] creates a binaural room impulse response (BRIR), which can be used for auralizations [82, 177].

RAVEN simulation software

Room Acoustics for Virtual ENvironments (RAVEN) [153] is a GA-based room acoustic simulation environment implemented in the C++ programming language. A short overview of the software and its development is presented as it was also used by one participant of the round robin presented in Chapter 4 and was applied by the author for all simulations in Chapter 5.

The RAVEN project was initiated at the Institute of Technical Acoustics at RWTH Aachen University by Dirk Schröder in 2005 and is mainly documented in his PhD thesis [151]. Its simulation core is based on a hybrid simulation model combing the image source model and a ray tracing algorithm, which were developed and elaborated in earlier research projects at the same institute [174, 86, 89].

The extension of the ray tracing algorithm with the diffuse rain technique [66] that calculates the amount of scattered energy projected to the receiver for each wall reflection, can also be regarded as a first order radiosity implementation. Furthermore, the RAVEN simulation model was extended by a diffraction module [152] and by a building acoustics module [182]. Both of these modules, however, require a special configuration of simulation and input data and are typically not activated by default.
Advantages of RAVEN are the efficiency of the simulation and the filter synthesis of BRIRs, which makes it possible to create convincing binaural auralization in real-time, e.g., in virtual acoustic environments [92]. This is achieved by spatial data structures [150] and shared memory multiprocessing programming using OpenMP.
Since 2011, various modules have been added to RAVEN. Sönke Pelzer, the second main contributor to this software, integrated a level-of-detail approach [127] and added options to generate RIRs for spatial reproduction techniques such as vector base amplitude panning (VBAP) and higher-order ambisonics (HOA) [124]. Furthermore, the MATLAB interface of RAVEN was extensively extended, allowing script-based configuration and operation of simulations, which is attractive for researchers in room acoustics and virtual acoustics. The introduction of an interface with the 3D modeling software SketchUp allows real-time visualization of parameters [123] and auralization [15] while the user is able to modify the virtual scene inside the SketchUp tool. Today, the RAVEN software is used for research by various groups and in university courses about room acoustics on five continents[1].

2.2.2 Validation of GA models

While GA models can be implemented efficiently in software [177, 151, 169], the limitations and assumptions of the ray-based sound propagation modeling should always be considered. The IS model itself, for example, ideally requires an infinite reflecting surface. ray tracing (RT) approaches have limited spatial and temporal resolution, and in general, simulations provide results only in the energy domain, based on random-incidence absorption coefficients, which are only valid in diffuse sound fields. Additionally, RT implementations include stochastic processes, which introduce further uncertainties. Neglecting wave-based effects becomes especially evident for low frequencies, where room modes dominate the sound field. This lower frequency limit of GA models is typically described by the Schroeder frequency [156].

[1] A software installer including the RAVEN simulation environment is freely available for academic purposes. Please contact raven@akustik.rwth-aachen.de for details.

The validation of a GA-based simulation model is not straightforward. Ideally the result of the simulation has to be compared to an exact analytical solution of the wave equation of all, or at least typical room environments, for which the simulation is intended for. Defining an analytical solution would, however, only be possible for very simple room geometries such as a shoebox room with rigid surfaces. This result can be used as a reference for a GA simulation of the same scenario, but it would only show deviations of the GA model in a situation, where reflections are purely specular and thus would fail to validate implementations with respect to surface scattering, which is of relevance for every real room. Reference results of more complex rooms could be obtained by wave-based models such as FEM simulations, as it was done for a *joint cuboid* room in a study by Southern et al. [162], but the issue of realistic geometry boundaries remains. In any case, it is challenging to find adequate scenarios for which both analytical or wave-based models and GA models can apply identical input data with respect to geometry and boundary conditions.

For these reasons, the common validation approach is an empirical validation of selected scenarios, i.e., the comparison with room acoustic measurements. Example studies can be found in [172, 122, 101, 170]. Most studies compare room parameters, in some cases they also discuss deviations in time and frequency domain and occasionally, also a perceptual comparison using listening experiments [133] is conducted. It should additionally be noted, that also the measurement does not perfectly describe the real situation. In order to make simulations and measurements comparable, the characteristics of non-ideal sound sources and receivers must be accounted for, as it was done in a validation study by Tsingos et al. [172]. In this study, conducted for a small model room of 16 m^3, a sound source with a lower frequency limit at 100 Hz was selected. If the sound source had been replaced by a larger measurement speaker with a lower cut-off frequency, a relevant modeling issue for the corresponding simulation would have been introduced as the sound source would have been a relevant object in the rather small room.

In order to provide an extensive database for future validations and evaluations of room simulations, the next chapter describes the creation a collection of acoustical scenes including the relevant input data required for the simulation of these scenes.

3

Scene database

The lack of confidence in the accuracy of room acoustical simulations becomes apparent when it comes to the design of new performance venues for music and speech, where acoustic scale models are still an important alternative with specific advantages [139]. The wide range of application fields of acoustical simulation necessitates a comprehensive evaluation of the applied algorithms, especially if considering that all of them have underlying simplifying assumptions or a limited valid frequency range. This chapter presents a collection of scenes, which were chosen in order to provide a common database for the evaluation and validation of simulation algorithms.

Version history

The first version was created in 2016 to provide input data to the participants of the round robin investigation on room acoustic simulation and auralization which is presented in Chapter 4. As the participants should not be informed about the measurement results, at this stage, the database included only input data and not the corresponding measurement data. These measurements were added to the database in the first research data publication in 2018 [11]. This version of the database was initially called *Ground Truth for Room Acoustical Simulation* (GRAS), but was later renamed and published as the Benchmark for Room Acoustical Simulation (BRAS) [12]. Internal revisions and peer-reviews of related publications helped to correct mistakes and improve the documentation and extend the database, leading to the most recent version of the database [13]. An overview of the versions and the changes and additions to the database is given in Table 3.1. As the data is uploaded on research data repositories using digital object identifiers (DOIs), a long-term availability is guaranteed and users are able to compare the different versions of the database. There are no specific plans yet for new versions of the database, but in general extensions are possible, e.g., more scenes or improved datasets of the boundary conditions.

Table 3.1: Different versions of the database including corresponding lists of changes and additions

1	**Database for first phase of round robin project** (2016) - initial version including the scene descriptions and input data - sent to all participants of round robin project
2	**Ground Truth for Room Acoustical Simulation** **GRAS**, Version 1, (2018, [11]) - addition of reference scenes 6 and 7 (RS6 and RS7) - addition of reference measurement results for all scenes - correction of some scene descriptions
3	**Benchmark for Room Acoustical Simulation** **BRAS**, Version 2, (2019, [12]) - first major update: Renaming and revision of database - general update of documentation and surface descriptions - correction of minor mistakes in scene descriptions - added *fitted* absorption coefficients for complex rooms - added detailed photo documentation for complex rooms - added geometric details of 3D models in scenes 9, 10 and 11 - added more detailed description of dodecahedron speaker - added directivity data of dodecahedron speaker
4	**Benchmark for Room Acoustical Simulation** **BRAS**, Version 3, (2020, [13]) - general update of documentation - renamed scenes 1-7 to reference scenes 1–7 (RS1–7) - renamed scenes 8-11 to complex rooms 1–4 (CR1–4) - complex rooms 1–4 (CR1–4) moved to the appendix - added list of 3D model simplifications for complex rooms - updated surface description and scattering coefficients - correction of minor mistakes in scene descriptions

3.1 Concept

The BRAS database presented in this chapter contains in total eleven acoustical scenes. The reference scenes 1–7 (RS1–7) highlight certain acoustical phenomena and certain spatial configurations, so that they can be used as a reference to evaluate the ability of room acoustical simulation software to model these phenomena and configurations. In addition, the database contains four complex

environments of different size and shape (scenes 8–11, CR1–4). An overview of all scenes is given in Table 3.2. To evaluate corresponding simulations both in the physical and in the perceptual domain, impulse responses (IRs) have been measured using omnidirectional receivers (all eleven scenes) and binaural receivers (for seven scenes), in some cases for different scene configurations, e.g., different reflection surfaces. All scenes are presented in the following subsections, more details including dimensions, orientations and positions of all elements for all scene configurations are provided in the data publication of BRAS [12].

Table 3.2: Overview of all scenes: Scenes 1-7 are designed reference scenes (RS), scenes 8-11 are representative complex rooms (CR). Most scenes are measured for multiple source (src) and receiver (rec) positions, and for more than one configuration (cfg), but not necessarily for all all source/receiver combinations.

Scene	**RIRs (cfg/src/rec)**	**BRIRs (cfg/src/rec)**
RS1: single reflection (infinite plate)	27 (3/3/3)	3 (3/1/1)
RS2: single refl. & diffr. (finite plate)	18 (4/6/5)	-
RS3: multiple refl. (parallel plates)	1 (1/1/1)	1 (1/1/1)
RS4: single reflection (reflector array)	18 (2/3/3)	-
RS5: diffraction (infinite wedge)	16 (1/4/4)	1 (1/1/1)
RS6: diffraction (finite body)	9 (1/3/3)	-
RS7: multiple diffraction (seat dip)	8 (1/2/4)	-
CR1/8: coupled rooms	8 (2/2/2)	2 (1/2/2)
CR2/9: seminar room (small)	10 (1/2/5)	5 (1/5/1)
CR3/10: chamber music hall (medium)	10 (1/2/5)	5 (1/5/1)
CR4/11: auditorium (large)	10 (1/2/5)	5 (1/5/1)

3.2 Reference measurements

The single channel measurements of all scenes were conducted and post-processed using the ITA-Toolbox [25], which provides an acoustical measurement environment for the MATLAB software. For all measurements, exponential sweeps were selected as an excitation signal. The sweep-based measurement method obtains IRs by deconvolution of the output signal and the measurement signal [54, 103]. To cover the full audible frequency range from 20 Hz to 20 kHz, the sweep signals also covered this frequency range, independent from the operation range of the

used loudspeakers. The length of the sweep varied from scene to scene, but in most cases was set to $2^{18}/44100\,\mathrm{Hz} = 5.94\,\mathrm{s}$. In case of the single channel IRs, all measurements were conducted four times and then averaged. BRIRs were measured with the FABIAN head and torso simulator measurement system for Matlab [93].
Although the database is not intended for the validation of sound pressure levels and source (power) modeling, the signal input chain was calibrated with a microphone calibrator (B&K Type 4231) and the output chain was calibrated to a free field sound pressure of 80 dB in front of the loudspeaker (on-axis) at a distance of 2 m.
A list of the used equipment during the reference measurements and a description of the measurement environment including the temperature and relative humidity is provided in the database documentation [13].

Postprocessing

Postprocessing of the simple scene results presented in this chapter included the application of a two-sided time window. If not stated otherwise, the fade-in was set to $\Delta t_{\mathrm{in}} = 3\,\mathrm{ms}$, applied 3 ms before the arrival of the direct sound, the fade-out was set to $\Delta t_{\mathrm{out}} = 10\,\mathrm{ms}$ at a point in time after all relevant reflections have occurred. In case of time domain results, a high-pass filter using the lower threshold frequency of the corresponding anechoic measurement room. Therefore, for scenes 1, 5, 6 and 7 the high-pass was set to $f_{\mathrm{cuttoff}} = 100\,\mathrm{Hz}$ and for scenes 2, 3 and 4 it was set to $f_{\mathrm{cuttoff}} = 63\,\mathrm{Hz}$, using the *ita_mpb_filter* method of the ITA-Toolbox.
To calculate energy decay curves and room parameters based on the measured RIRs of the complex rooms, the ITA-Toolbox methods *ita_roomacoustics* and *ita_roomacoustics_EDC*, respectively, were applied to the RIRs.

3.3 Scene descriptions

Scene 1 (RS1): Single reflection (infinite plate)

Scene 1 features a single reflection for different angles of sound incidence on a quasi infinite surface. To achieve a plane wave reflection, the corresponding surfaces have been chosen as large as possible in order to treat them as infinite reflectors. The three configurations of the scene include different types of reflecting surfaces:

1. Rigid surface (floor of hemi anechoic chamber)
2. Absorbing surface (thickness: 20 mm)
3. diffuser (rectangular profiles)

To consider angle dependent reflection effects, three positions have been defined each for sound source and receiver which correspond to the angles of incidence $\theta = \{30°, 45°, 60°\}$. The side view of the scene for configuration 3 including the diffuser profiles made from medium-density fibreboard (MDF) is shown in Fig. 3.1. The distance of all sound source and receiver positions was set to

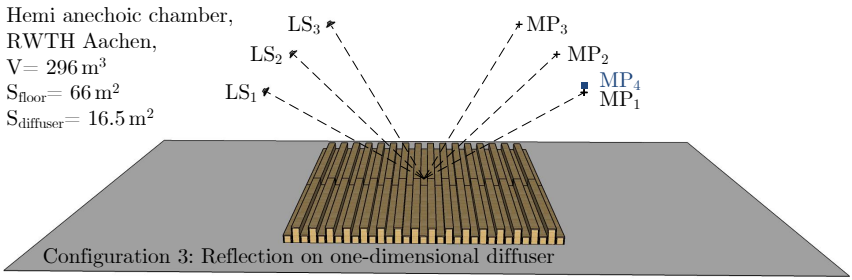


Figure 3.1: Side view of scene 1 for configuration 3 including three sound source positions and four receiver positions. The blue square (MP4) indicates the position of the binaural receiver.

3 m, for higher distances it would have been challenging to accurately position loudspeakers and microphones in the hemi anechoic chamber, which was chosen for the reference measurements. RIRs of configurations 1 – 3 have been measured for all combinations of LS1 – LS3 and MP1 – MP3. BRIRs of configurations 1 – 3 have only been measured for sound source LS2 and receiver position MP4. The finite size of the reflectors causes diffraction effects, which start to matter below a threshold frequency f_g, which is calculated according to Eq. (3.1) [137, 141]. Below this threshold frequency, a specular reflection is attenuated by 6 dB per octave.

$$f_g(\theta) = \frac{c \cdot a^*}{2 \cdot S_{refl} \cdot cos(\theta)}, \tag{3.1}$$

with S_{refl} being the surface area of the reflector, a^* the characteristic distance, c the speed of sound and θ the angle of incidence. As the distance from the reflection point to all sound source and receiver positions is identical $a = 3\,\text{m}$, the characteristic distance a^* is also 3 m. Table 3.3 includes the surface areas of the reflectors and the calculated values of f_g for all configurations and the angles $\theta = \{0°, 30°, 45°, 60°\}$. The normal incidence $\theta = 0°$ was included for reference. In configuration 1, the reflection occurs on the rigid floor of the hemi anechoic

chamber and has a substantially larger surface area than the absorber panels and the diffuser in configuration 2 and configuration 3, respectively. The highest threshold frequency in this evaluation is 62.3 Hz, which roughly corresponds to the lower frequency cut-off of the loudspeaker used in the reference measurements (cf. Section 3.4).

Table 3.3: Surface areas of reflector and threshold frequencies f_g for scene configurations 1–3 for different angles of incidence

	S_{refl}	$f_g(0°)$	$f_g(30°)$	$f_g(45°)$	$f_g(60°)$
configuration 1 (rigid floor)	65.67 m^2	7.8 Hz	9.0 Hz	11.1 Hz	15.7 Hz
configuration 2 (absorber)	17.64 m^2	29.2 Hz	33.7 Hz	41.2 Hz	58.3 Hz
configuration 3 (diffuser)	16.53 m^2	31.1 Hz	35.9 Hz	44.0 Hz	62.3 Hz

Reference measurements

Reference measurements of IRs of all three configurations for loudspeaker position LS3 and MP3 are shown in Fig. 3.2. In the graph, two impulse are visible, leading to a typical comb-filter structure in the frequency domain. The direct sound, which coincides for the three configurations, has the peak at 8.7 ms, which corresponds to the distance of 3 m between sound source and microphone. The reflection arrives after 17.5 ms, corresponding to the propagation distance of 6 m of the reflected wave. The highest amplitude of the reflection can be observed for the rigid floor, as expected. Reflections on the absorber and the diffuser are attenuated by more than 5 dB. Additionally, the diffuser shows substantially more reflected energy up to 24 ms than the other two configurations.

Scene 2 (RS2): Single reflection & diffraction (finite plate)

Scene 2 has a similar setup as scene 1, the reflecting panel however, is substantially smaller. A sketch of the scene including the sound source and receiver positions is shown in Fig. 3.3. No binaural, but only omnidirectional receivers are considered in this scene. Sound sources and receivers are arranged similarly to scene 1 for three angles of sound incidence, $\theta = \{30°, 45°, 60°\}$. Reflections on the panels can be analyzed for sound source positions LS1 – LS3 and receiver positions MP1 – MP3. By adding two receiver positions behind the panel (MP4 and MP5), the

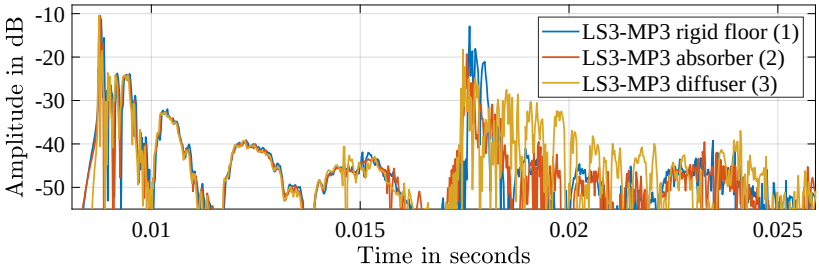


Figure 3.2: Scene 1 reference measurement of all three configurations for sound source position LS3 and receiver position MP3

scene also allows investigations of diffraction and sound transmission. Despite its simplicity and in contrast to scene 1, this scene is challenging especially for GA-based simulations as diffraction and sound transmission around and through the plate have to be considered in the applied simulation software. The scene

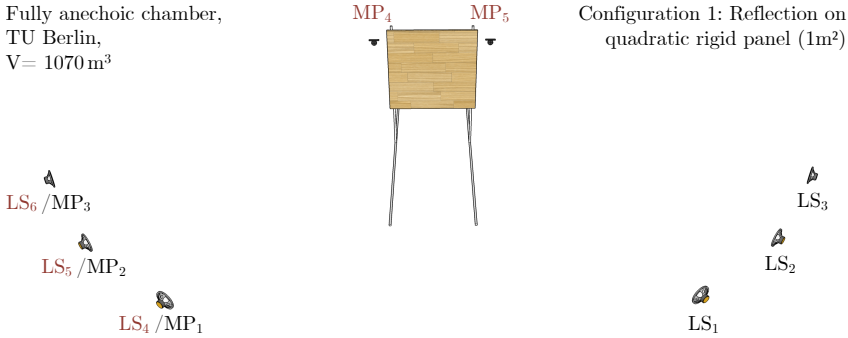


Figure 3.3: Scene 2 for configuration 1 including six sound source positions and five receiver positions. The panel is mounted on a stand which was placed on the wire floor of the fully anechoic chamber.

is investigated for four different configurations, varying in size and type of the quadratic reflector panel:

1. Rigid panel (edge length: 1 m, thickness: 25 mm)
2. Absorber panel (edge length: 1 m, thickness: 45 mm)
3. Rigid panel (edge length: 2 m, thickness: 25 mm)
4. Absorber panel (edge length: 2 m, thickness: 45 mm)

The threshold frequency of this scene, calculated according to Eq. (3.1) for the characteristic distance $a^* = 4\,\mathrm{m}$, are shown in Table 3.4. In comparison to scene 1, the calculated threshold frequencies are considerably higher, especially for the smaller reflection panel with a size of $1\,\mathrm{m}^2$, attenuated reflections are to be expected even for frequencies above 1 kHz in case of a sound incidence of $\theta = 60°$.

Table 3.4: Surface areas of reflecting panels and threshold frequencies f_g for scene configurations 1–4 for different angles of incidence

	S_{refl}	$f_g(30°)$	$f_g(45°)$	$f_g(60°)$
configuration 1 & 2	$1\,\mathrm{m}^2$	792 Hz	970 Hz	1372 Hz
configuration 3 & 4	$4\,\mathrm{m}^2$	198 Hz	243 Hz	343 Hz

Reference measurements

Reference measurements were conducted for the four configurations for all combinations (9) of sound sources LS1–LS3 and for receiver positions MP1–MP3, IRs for all combinations (6) of sound sources LS4–LS6 and for receiver positions MP4–MP5 were measured for configuration 1–3, resulting in a total number of 54 reference IRs for this scene. Results of all four configurations in the frequency domain for loudspeaker position LS2 and MP2 (specular reflection for $\theta = 45°$) are shown in Fig. 3.4. It can be observed that for frequencies lower than 400 Hz, the results of configuration 1 and 2, and, the results for configuration 3 and 4 coincide. In this frequency range, only the geometrical aspects of the reflecting panel are relevant as both, the absorber and the MDF panel hardly absorb any incident energy. For frequencies above 1 kHz, curves for configuration 2 and 4 coincide, and also the curves of configuration 1 and 3 show resembling behavior. In this frequency range, the specular reflection is not influenced by the finite size of the panel anymore (cf. threshold frequencies in Table 3.4) and thus only depends on the surface material. The curves of both absorbing panels show a weaker comb filter structure as more energy of the reflected wave is absorbed for increasing frequencies.

Scene 3 (RS3): Multiple reflection (parallel finite plates)

Scenes 3 aims at recreating another simplified versions of a relevant real life scenario: The reflection between parallel plates, which evokes a flutter echo that is often problematic in room acoustics. The sketch of the scene, which only includes

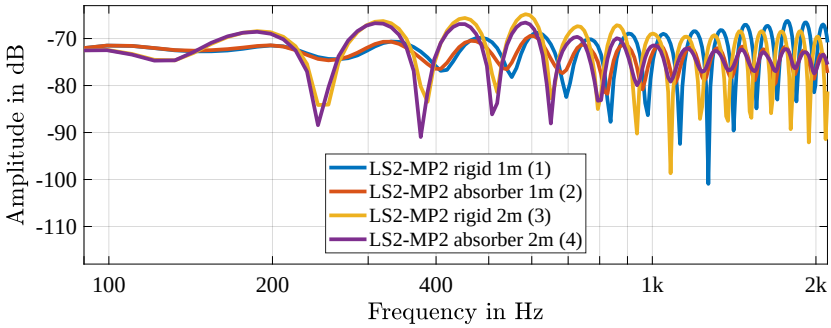


Figure 3.4: Scene 2 reference measurement of all four configurations for sound source position LS2 and receiver position MP2 in the frequency domain

one sound source and one receiver, is shown in Fig. 3.5. Receiver and sound source are positioned on the center axis between parallel MDF panels of $4\,\mathrm{m}^2$ in an otherwise anechoic environment. The scene is defined for an omnidirectional receiver (MP1) and for a binaural receiver (MP2), with the interaural axis coinciding with the center axis and the right ear facing the sound source position LS1. In this scene, the threshold frequency f_g calculated according to Eq. (3.1) of a normal incidence reflection ($\theta = 0°$) for a characteristic distance $a^* = 4.2\,\mathrm{m}$ is 180 Hz.

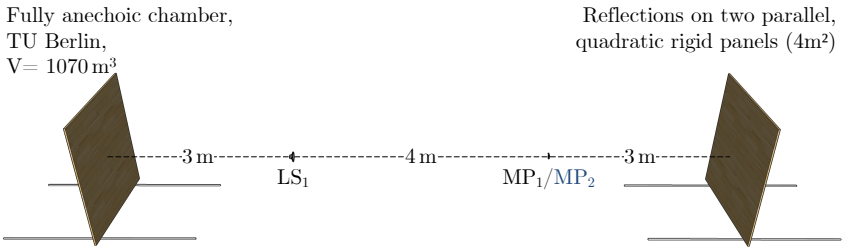


Figure 3.5: Sketch of scene 3 including source position LS1, one omnidirectional (MP1) and one binaural receiver position (MP2). The parallel plates are located at a distance of 10 m from each other.

Reference measurements

In Fig. 3.6, the resulting IR reference measurement for loudspeaker position LS1 and MP1 are shown. The simple geometry of the scene allows a straightforward calculation of image sources in order to model the reoccurring reflections caused by the parallel panels. In the graph, image sources of an omnidirectional point source are displayed up to the 7th order in addition to the measured results. The source power of the sound source has been adjusted to match direct sound levels of measurements and model. The results of the model temporally coincide with the direct sound and reflection peaks of the measured IR and thus validate the given dimensions of the scene.

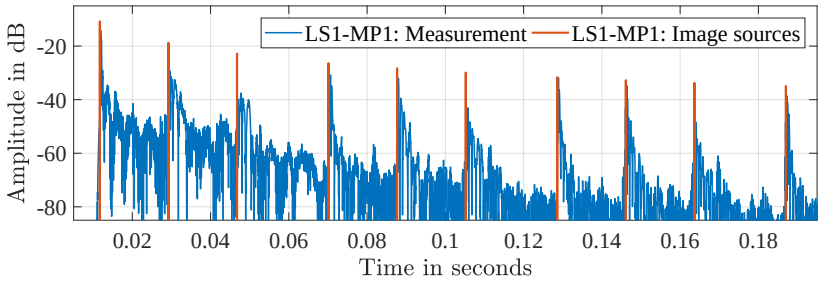


Figure 3.6: Scene 3 reference measurement and image sources for sound source position LS1 and receiver position MP1 in the time domain

Scene 4 (RS4): Single reflection (reflector array)

Scene 4 uses a very similar setup as scene 2, and is sketched in Fig. 3.7. Instead of one reflector panel, an array consisting of nine quadratic elements with an edge length of 0.68 m ($S_{refl} = 0.46\,\mathrm{m}^2$) is situated in an anechoic environment. This scene is included in the database as reflector arrays are frequently used in concert halls to direct early reflections to the audience area. To investigate different incidence angles $\theta = \{30°, 45°, 60°\}$, three sound source and three receiver positions are defined, all at a distance of 3.80 m from the center of the reflector array. Two configurations are defined for the scene:

1. Reflection point located on center of panel array (on-center)
2. Reflection point located in the gap between four panels (off-center)

The sketch for configuration 2 is depicted in Fig. A.1 in Appendix A.1.
According to the theory of Rindel [137], the lower limiting frequency of the array

with a total surface area $S_{total} = 5.34\,\mathrm{m}^2$ can be calculated using Equation (3.2):

$$f_{g,total}(\theta) = \frac{c \cdot a^*}{2 \cdot S_{total} \cdot cos(\theta)}, \tag{3.2}$$

In case of the reflection point being on the panel (configuration 1), above this frequency, and up to $\mu \cdot f_g$, the level of the reflection is attenuated by:

$$\Delta L_{\mathrm{diffr}} \simeq 20\ \log(\mu) = 2.2\,\mathrm{dB} \tag{3.3}$$

with $\mu = 9 * S_{refl}/S_{total} = 0.78$. For frequencies higher than f_g, no attenuation for reflections is observed as long as the reflection takes place on an individual panel of the array. The calculated frequencies for the three sound sound incidence angles are listed in Table 3.5.

Table 3.5: Surface areas of reflecting panels and threshold frequencies f_g for different angles of incidence.

	$\theta = 30°$	$\theta = 45°$	$\theta = 60°$
$f_{g,total}$	141 Hz	173 Hz	244 Hz
$\mu \cdot f_g$	1269 Hz	1554 Hz	2198 Hz
f_g	1627 Hz	1993 Hz	2819 Hz

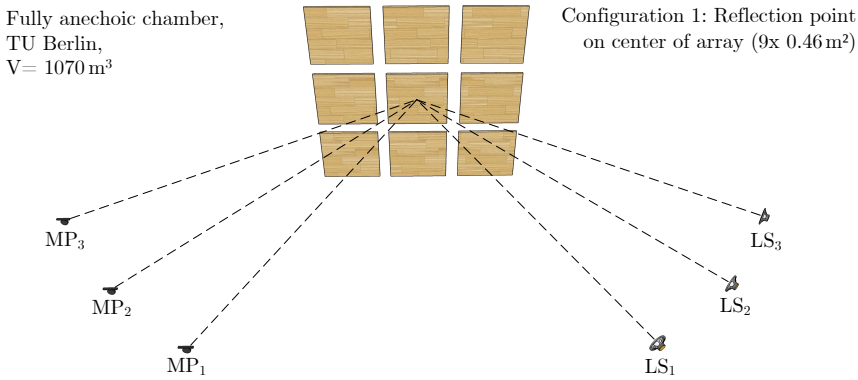


Figure 3.7: Side view of scene 4 for configuration 1 including three sound source positions and three receiver positions

Reference measurements

Reference measurements of IRs for LS2–MP2 (configuration 1), and for LS5–MP5 (configuration 2) are shown in Fig. 3.8. While the direct sound for both configuration coincides as expected, the peak level of the reflection is attenuated by more than 7 dB. For configuration 2, less reflected energy is mostly observed for higher frequencies, as the reflection is primarily attenuated when the wavelength is substantially smaller than the gap of around 13 cm. In Fig. 3.9 the frequency response of the reflection on the reflector array is depicted. For this, the reflection was isolated using a two-sided hann window for $t_{\text{hann2}} = \{21, 22, 25, 26\}$ ms. The depicted curves were normalized and contain the frequency response of the loudspeaker. On axis, this frequency response is, however, almost flat for the selected frequency range (cf. Section 3.4). The measured frequency response of the reflections include variations of more than 10 dB and do not resemble the predicted behavior of the model for both configurations. Nevertheless, the deviations are mostly lower than 5 dB for frequencies below 500 Hz, and in case of the reflection point being in the gap (configuration 2), for the entire depicted frequency range. For high frequencies above 2 kHz, in both configurations the characteristic trend of the model can also be observed. While for configuration 1, the energy of the reflection remains fairly constant above 2 kHz, the energy for configuration 2 decays with increasing frequency, and resembles, on average, the predicted 6 dB per octave decay.

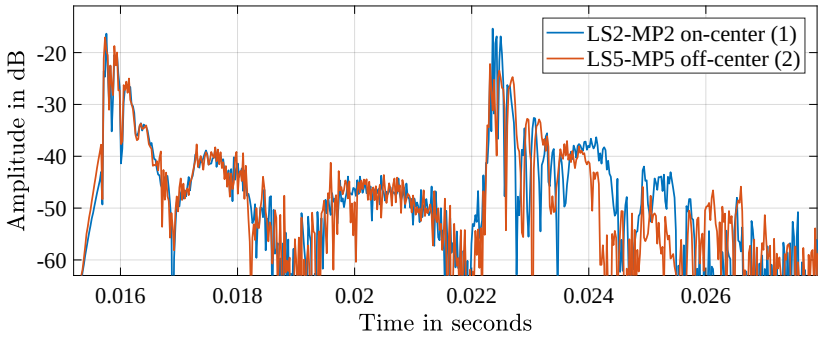


Figure 3.8: Scene 4 reference measurement of both scene configurations for impulse responses LS2-MP2 and LS5-MP5

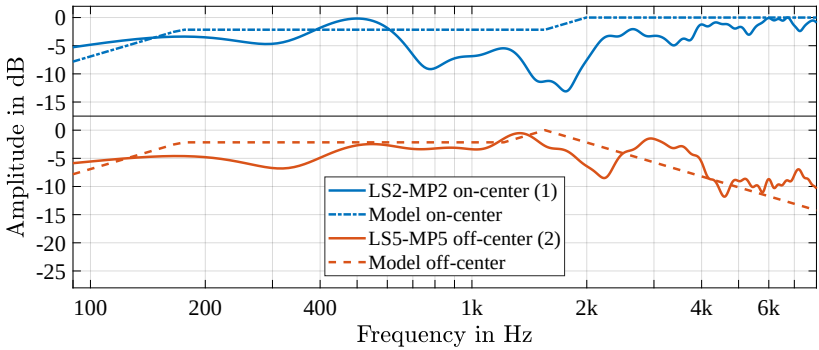


Figure 3.9: Frequency responses of reflections on reflector panel for positions LS2-MP2 and LS5-MP5 including the attenuation calculated according to Rindel [137]. Corresponding values for $f_{g,total}$, $\mu \cdot f_g$ and f_g can be found in Table 3.5 ($\theta = 45°$).

Scene 5 (RS5): Diffraction (infinite wedge)

Scene 5 is situated in a hemi anechoic environment and includes a wall, made from MDF with a thickness of 25 mm, which separates the four sound source from the four omnidirectional receiver positions, see Fig. 3.10. A binaural receiver, located at receiver position MP1, was also considered for sound source position LS1. The MDF wall of the scene represents a diffraction wedge, which is especially relevant for the acoustics of urban environments. With the mass per area $m'' = 18.56\,\mathrm{kg/m^2}$, the sound reduction index R according to the mass law of a single-layer wall at f =100 Hz is 23 dB, calculated with Eq. (3.4):

$$R \approx 20 \log \frac{2\pi f \cdot m''}{2\rho_0 c}\ \mathrm{dB} \tag{3.4}$$

with c being the speed of sound and ρ_c the density of air. This result shows that a substantial amount of energy is not transmitted through the wedge and that thus diffraction effects dominate the sound energy at the receiver positions without a direct line of sight to the sound source position.

Reference measurements

Reference measurements in the frequency domain for sound source position LS2 and three receiver positions are shown in Fig. 3.11. A comb filter structure for the three positions is visible especially for frequencies below 500 Hz. This effect

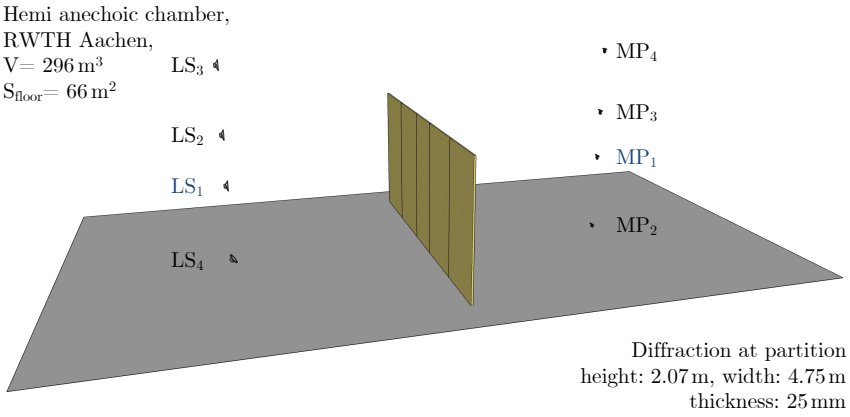


Figure 3.10: Side view of scene 5 including four sound source and four receiver positions in a hemi anechoic environment. For sound source position LS1, MP1 is defined as both, an omnidirectional and a binaural receiver.

is caused by reflected sound waves, either before or after being diffracted on the partition. For combination LS2-MP4 the direct path between sound source and receiver is not occluded by the partition leading to a rather flat frequency response for frequencies above 1 kHz. The comb filter effect is nearly vanished in this frequency range as the most relevant diffraction only occurs for lower frequencies. This can also be observed for the other receiver positions, which show substantially less energy for higher frequencies, especially when the receiver is positioned on the floor (MP4).

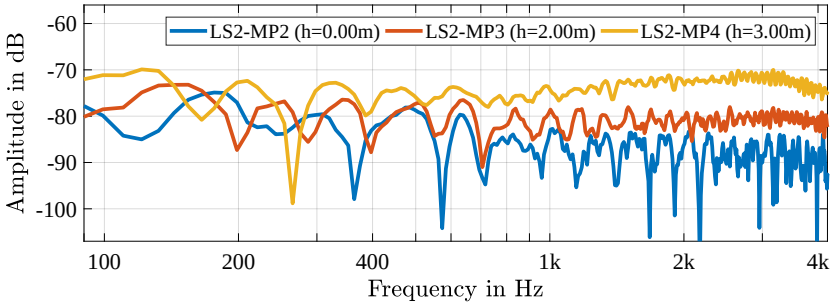


Figure 3.11: Scene 5 reference measurement for sound source position LS2 and three receiver positions MP2-MP4. The variable h indicates the height of the receiver.

Scene 6 (RS6): Diffraction (finite body)

Scene 6 is similar to the previous scene and also situated in a hemi anechoic environment. Instead of a partition, a cubic diffraction body of the same width is placed between the three sound sources and the three receivers, which are located at a distance of 3 m in front and behind the cubic body, respectively. The body has a height and depth of 72 cm, two sound source and two receiver positions are located at a lower level than the body, while the third sound source and the third receiver are both positioned at a height of 80 cm. This leads to a direct line of sight only between LS3 and MP3. A view from the side is given in Fig. 3.12. The scene does not correspond directly to a typical real-world scenario, but was designed as a scale model, i.e., 1:6 or 1:10, of a structure in an urban environment, which typically has a relevant depth and thus includes multiple diffraction edges.

Reference measurements

Reference measurements in the frequency domain of sound source position LS2 and two receiver positions, MP2 and MP3, are shown in Fig. 3.13. Below 1 kHz both curves exhibit an irregular comb filter effect, created by diffracted sound waves at the top and the sides, either directly arriving at the receiver after being diffracted or reflected on the floor. Notches are shifted by frequency and varying in level because of the different height of two selected receiver positions. For higher frequencies, diffraction only remains on the top of the body, creating a more regular comb filter structure.

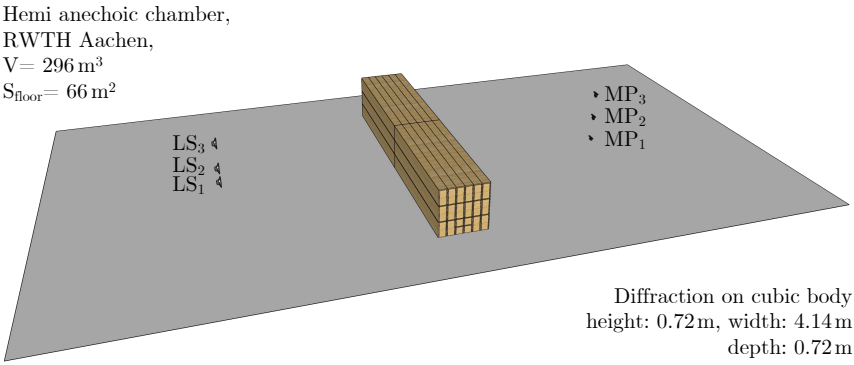


Figure 3.12: Side view of scene 6 including three sound source and three receiver positions in a hemi anechoic environment

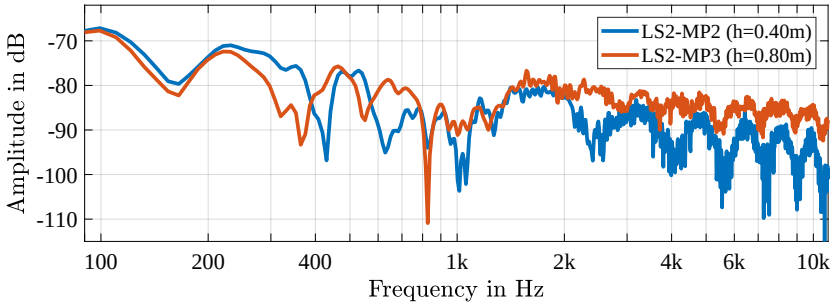


Figure 3.13: Scene 6 reference measurement in the frequency domain for sound source position LS2 and receiver positions MP2 and MP3

Scene 7 (RS7): Multiple diffraction (seat dip effect)

In concert halls, the seat-dip effect [157] describes an additional attenuation at low frequencies when the sound is traveling from the stage to the end of the audience area. Especially for GA-based room simulations, it is very challenging to account for this effect in the simulation model. For this reason, scene 7 models seating rows using 15 rectangular profiles, separated by a gap of 22 cm, with a height of 24 cm and a depth of 12 cm, each made of MDF. The arrangement of the 15 profiles in the hemi anechoic environment including all sound source and receiver positions is depicted in Fig. 3.14. One sound source position is located slightly

below the height of the rows at a height of 18 cm, the second source is positioned at a height of 52 cm, tilted by 10° downwards. The distance of both sound sources to the first row is 77 cm. Two receiver positions are located in the center at a height of 34 cm and 40 cm, between row 7 and row 8, two more receivers are positioned behind the last row at the same heights. These positions were selected to represent a typical ear height of a seated and a standing person, respectively, considering a scale of 1:4. According to the theory of the seat-dip effect [35], for a row height of 24 cm, the frequency of the dip is $f_{seatdip} = \frac{c}{4 \cdot h} \approx 357$ Hz.

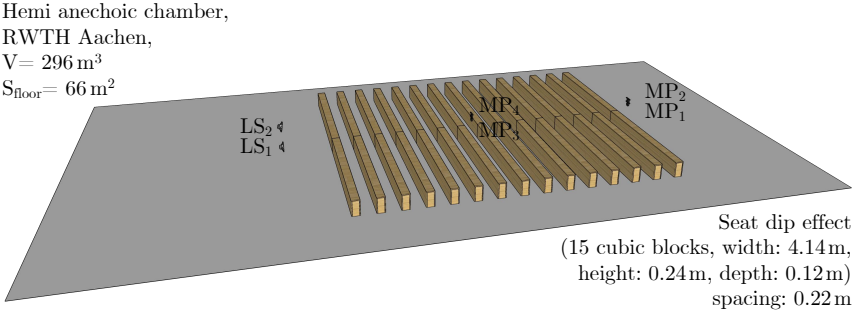


Figure 3.14: Side view of scene 7 including two sound source and four receiver positions in a hemi anechoic environment

Reference measurements

Reference measurements in the frequency domain of sound source position LS2 and receiver positions MP1 (behind the array of profiles) and MP3 (in the center of the array) are shown in Fig. 3.15. The measured IRs were time windowed with a two-sided *hann* window with $t_{\text{hann2}} = t_{\text{direct}} + \{-6, -3, 10, 15\}$ ms, and the results in the frequency domain were smoothed using a one-third octave band moving average filter. When comparing the results of both microphones, the most prominent difference is the lower level of around 6 dB for receiver position MP1. This attenuation corresponds to the spherical spreading loss due to the larger distance Δd to the sound source for receiver MP1. While the seat dip effect in concert halls is typically measured between 100 Hz and 500 Hz [157], for the scaled setup, the seat dip is shifted towards higher frequencies, and lies, in both cases, between 350 Hz and 400 Hz. This matches the theoretical dip frequency of 357 Hz at which the seat height corresponds to a quarter of the wavelength. The results also show that the dip frequency is higher for MP1, indicating that the

dip frequency can be shifted towards lower frequency if the angle of incidence is increased, as described in detail by Bradley [35].

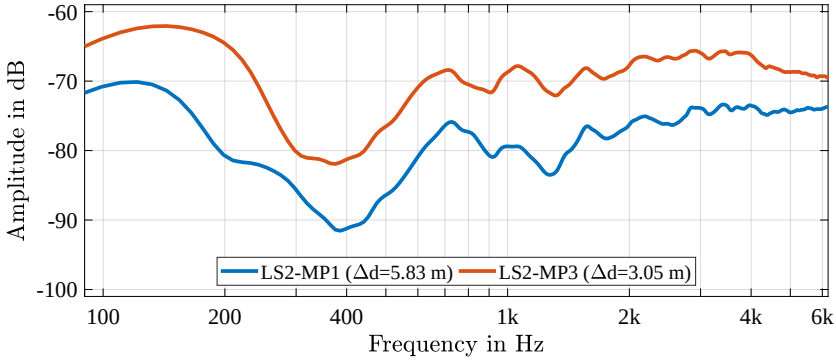


Figure 3.15: Scene 7 reference measurement in the frequency domain for sound source position LS2 and receiver positions MP1 and MP3. Results were smoothed using a one-third octave band moving average filter.

Scene 8 (CR1): Coupled rooms (laboratory room & reverberation chamber)

In this scene, two rooms with similar volume are connected, a reverberation chamber (R_2, $V_2 = 122$ m^3) and a significantly less reverberant laboratory room (R_1, $V_1 = 104$ m^3). The scene is investigated for two different opening angles of a rotating door, $\phi_A = 4.1°$ and $\phi_B = 30.4°$, leading to coupling areas (aperture size) of $S_{12}(\phi_A) = 0.12$ m^2 and $S_{12}(\phi_B) = 1.95$ m^2. The top view of the scene for ϕ_B is shown in Fig. 3.16. The receiver positions MP3 and MP4 are defined as omnidirectional receivers for both sound source positions. Additionally, two binaural receiver positions MP1, oriented towards LS1, and MP2, oriented towards LS2 are defined. Both are located at the same position as MP3. Binaural reference measurements were only measured for one sound source position and only for the door opening angle ϕ_B.
The two door opening angles were chosen to create two energy decays at different transition times. A photo of the door for ϕ_B, taken in R_1, is shown in Fig. 3.17a, in Fig. 3.17b it is explained how the aperture size is calculated from the corresponding 3D model (cf. [99]). The 3D model consists of 99 faces to which four different surface materials have been applied.

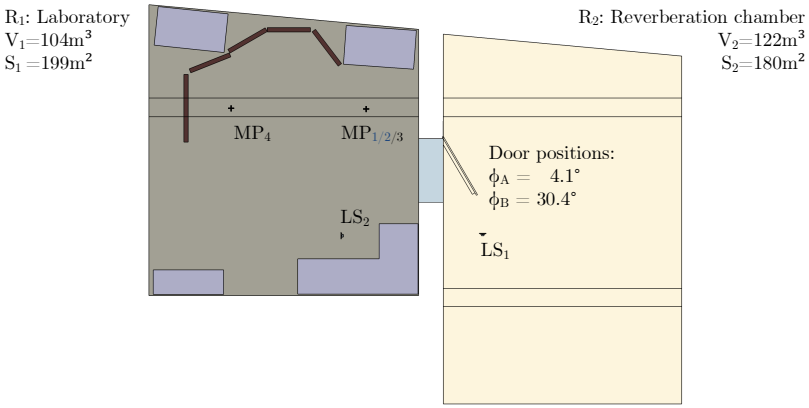
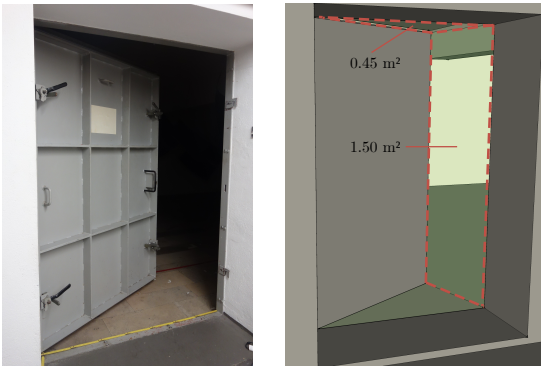


Figure 3.16: Top view of scene 8 for $\phi_B = 30.4°$. The laboratory room (R_1, left), is connected to a reverberation chamber (R_2, right) via a rotating door opened at different angles.



(a) (b)

Figure 3.17: Door for the opening angle ϕ_B during the reference measurements (a) and the corresponding 3D model of the scene (b) including the coupling areas resulting in a total aperture area of 1.95 m^2 for ϕ_B. As the door is mounted flush with the floor, but not with the ceiling, only the triangular surface above the door is considered in the calculation.

Reference measurements: Individual rooms

In both rooms, separate RIR measurements have been conducted while the door was closed ($\phi = 0°$). In the reverberation chamber, four RIRs have been measured for one source position, in the laboratory room in total ten RIRs have been measured, for two source positions and five receiver positions. The averaged reverberation times for R_1 and R_2 are $T20_{R1} = 0.78$ s and $T20_{R2} = 5.79$ s, respectively, yielding equivalent absorption areas of $A_{R1} = 21.49$ m^2 and $A_{R2} = 3.40$ m^2. The reverberation time ratio of both rooms can then be described by Eq. (3.5).

$$r_{meas} = \frac{T20_{R2}}{T20_{R1}} = 7.43 \tag{3.5}$$

Reference measurements: Energy decay curves of coupled rooms

For both opening angles $\phi_A = 4.1°$ and $\phi_B = 30.4°$, the energy decay curve (EDC) was evaluated for LS2 and MP3. In Fig. 3.18 both EDCs evaluated for the 1 kHz frequency band are shown. These results show a double slope for condition ϕ_A and ϕ_B and that a smaller opening angle of the door leads to lower late decay rate, while the early decay rate marginally changes. The transition from the early slope to the late slope can be either described by the *bent* level [88], the *bending*

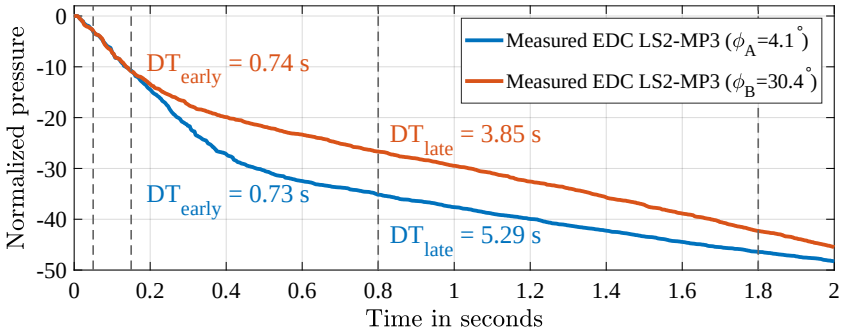


Figure 3.18: Measured energy decay curve (EDC) for both door opening angles $\phi_A = 4.1°$ and $\phi_B = 30.4°$, evaluated for the 1 kHz octave band for LS2 and MP3, both located in laboratory room. Dashed lines indicate the limits of the time frames used for the decay time evaluation.

point [98] or the *turning point* [189], which occurs at a later time at a lower level for ϕ_A in comparison to ϕ_B. To determine the decay rate β of the EDC in the early and the late part, a linear regression model (least-squares fit) is applied . The decay times (DT) are calculated for the early and the late part according to Eq. (3.6).

$$\text{DT [s]} = \frac{-60 \text{ dB}}{\beta} \tag{3.6}$$

The time windows to calculate DT_{early} and DT_{late} are [0.05 s .. 0.15 s] and [0.8 s .. 1.8 s], respectively, indicated by the dashed vertical lines in Fig. 3.18. These measures are chosen instead of alternative proposed measures, such as the EDT and the late decay time (LDT) [34]. The LDT, evaluating the decay time between the values of -25 dB and -35 dB, is not applied as the measured results contained non-linear decays between these levels. Choosing a time frame which includes a linear decay in case of all investigated scenarios thus represents a more suitable evaluation.

Scene 9 (CR2): Small room (seminar room)

The first room scene of the database contains a small empty seminar room at RWTH Aachen University. It was chosen because of its geometric simplicity and its rather small volume of 146 m^2. The absence of furniture of the room reduces an uncertainty factor when creating geometry for room simulations. The

small volume, however, makes this room challenging for pure geometrical acoustic based simulations. Two sound source and five omnidirectional receiver positions were defined corresponding to the procedure of ISO-3382 [76]. For the analysis of BRIRs, a setup containing five sound sources and one binaural receiver was defined. The omnidirectional receiver setup is shown in Fig. 3.19, the setup used for the binaural simulations and measurements can be found in Appendix A.1. The omnidirectional receiver positions were analyzed for both, an omnidirectional dodecahedron speaker and a near-field monitor, details on the sound sources are given in Section 3.4.

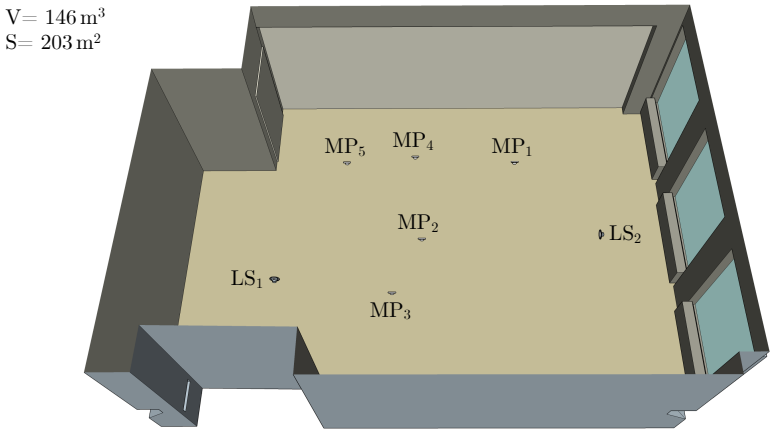


Figure 3.19: Top view of scene 9, an empty seminar room, for the setup including two omnidirectional sound sources and five omnidirectional receivers

Reference measurements

In addition to an analysis of the measured RIRs in the time and frequency domain, the measurements also allow an extensive analysis of various room parameters according to ISO-3382 [76]. As an example, in this chapter, only the most popular room parameter, the reverberation time, is evaluated and presented. In Fig. 3.20, the parameters EDT and T30 are shown, evaluated for six octave bands and averaged for all ten RIRs of the scene. The values confirm the expected high reverberation time, considering the rather small volume of the room. Interestingly the T30 value of 1.35 s for the 125 Hz octave band is lower than the corresponding value of the 4 kHz octave band of 1.55 s. This relatively high value for 4 kHz is

caused by the absence of porous absorber material in the room and and the small room volume, only causing minor air absorption effects.

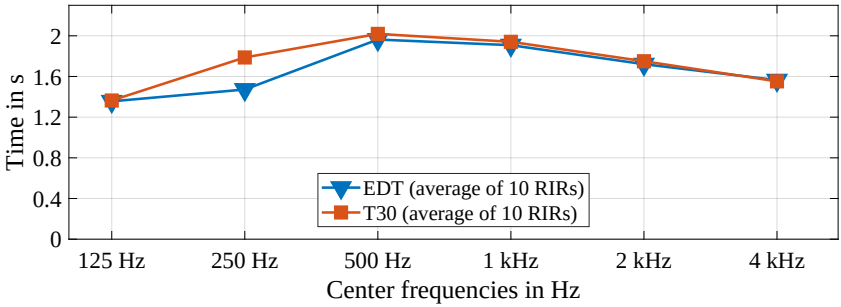


Figure 3.20: Measured reverberation times of scene 9. Values are averaged for two sound source and five receiver positions.

Scene 10 (CR3): Medium room (chamber music hall)

As a medium-sized room, the chamber music hall of the *Konzerthaus Berlin* was chosen. This prestigious hall for classical chamber music can be described as a shoebox room, it has, however, a coupled volume of almost $1000\,\mathrm{m}^3$ behind the stage and above the concert hall, which raises the question if and how a simulation software should account for this coupled volume. For this reason, the geometric model of the database includes the coupled space and all relevant apertures (see Fig. A.3, included in Appendix A.1). The stage of the hall can be varied in size, in this database, the extended stage configuration was chosen. For this setup, the stage has an approximate surface area of $80\,\mathrm{m}^2$, the main hall's audience area covers approximately $200\,\mathrm{m}^2$, including 238 seats[1], the gallery contains additional 104 seats. In an identical procedure to scene 9, this room is discussed for an omnidirectional and a binaural receiver setup. The two sound source and five omnidirectional receiver positions are shown in Fig. 3.21, a top view of the binaural setup is included in Appendix A.1. The values for the room volume and surface area, given in Fig. 3.21, correspond to the total volume and surface area of the provided 3D model, the corresponding values of the main hall only are substantially lower.

[1] Usually the parquet includes 250 seats, during the measurements, 12 seats were removed in the center of the audience area (around receiver position MP1)

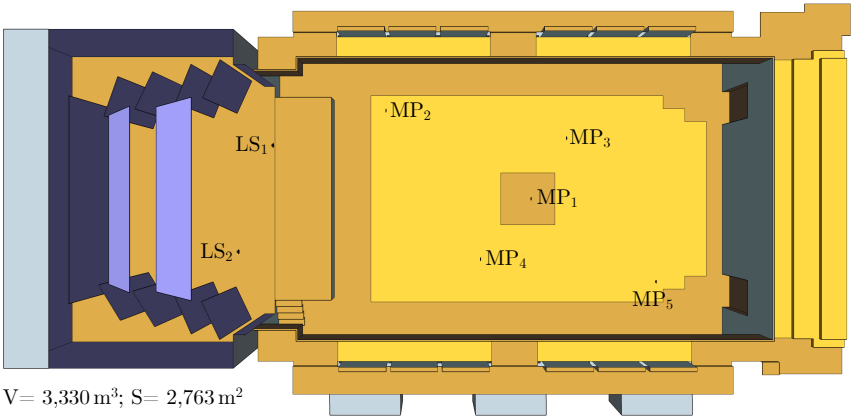


Figure 3.21: Top view of scene 10, the chamber music hall of the *Konzerthaus Berlin*, for the setup including two sound sources and five omnidirectional receivers

Reference measurements

Averaged values of the ten RIR measurements for the two reverberation time parameters EDT and T30 were obtained using the same procedure as for scene 9. The results for six octave frequency bands are shown in in Fig. 3.22. Although

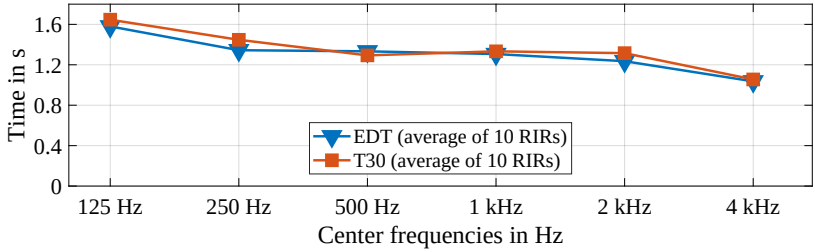


Figure 3.22: Measured reverberation times of scene 10. Values are averaged for two sound source and five receiver positions for six octave bands.

the room volume is much higher in comparison to the seminar room, the chamber music hall has a lower reverberation time for all frequency bands except for the 125 Hz frequency band. In general, the chamber music hall has a lower reverberation time than most typical chamber music halls (1.65 - 1.75 s), but lies

within the rather wide recommended range between 1.0 s and 2.0 s [23]. A large amount of the room's absorption can be attributed to the seating, consisting of upholstered wooden chairs.

Scene 11 (CR4): Large room (auditorium)

As a large room, the auditorium maximum of the Technical University in Berlin was selected. In contrast to scene 9 and 10, this lecture hall has a fan shape and contains reflectors, a balcony and a more complex ceiling structure. The volume of the provided 3D model is $8,657\,\mathrm{m}^3$, the surface area is $5,851\,\mathrm{m}^2$. In total, up to 1192 persons are seated on lightly upholstered wooden folding seats, located on the pitched floor (888) and the balcony (304) of the lecture hall. The room is is primarily used for lectures, but also for events and concerts of different musical styles. Its top view including the two sound source and five omnidirectional receiver positions is given in Fig. 3.23, the binaural setup is presented in Appendix A.1.

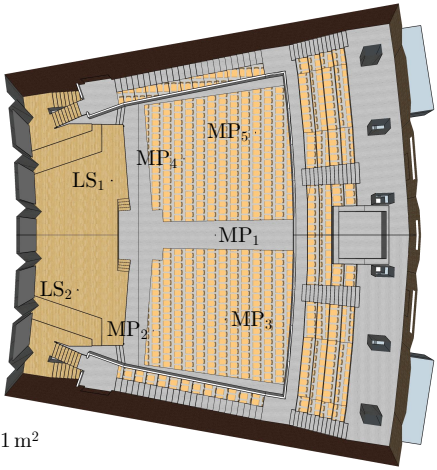


Figure 3.23: Top view of scene 11 (lecture hall at TU Berlin) for the setup including two sound sources and five omnidirectional receivers

Reference measurements

Analog to the other two single room scenes, the room parameters EDT and T30 are also presented for scene 11. The results for six octave bands between 125 Hz

and 4 kHz are shown in Fig. 3.24. The highest reverberation times of around 2.4 s are observed in the two lowest frequency bands, for 125 Hz and 250 Hz. For higher frequency bands, the reverberation times decrease to a value of 1.4 s for 4 kHz as the room's surface absorption and air absorption increases.

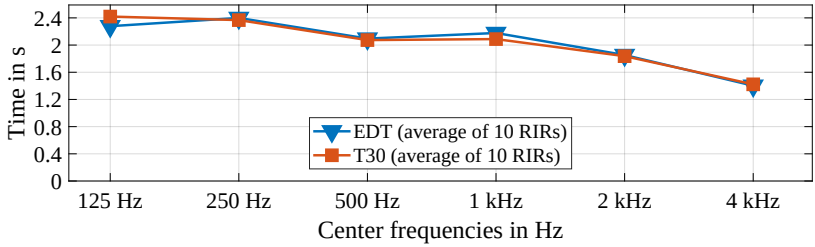


Figure 3.24: Measured reverberation times of scene 11. Values are averaged for two sound source and five receiver positions for six octave bands.

3.4 Sound sources

Depending on the purpose of the scene and its corresponding measurement, different sound sources are selected. The three sound sources considered in the BRAS database are depicted in Fig. 3.25. For the simple scenes of the database

Figure 3.25: Three sound sources of the BRAS. Red crosses indicate the reference points of each source, red dots indicate the reference points of the subwoofer and the tweeter of the dodecahedron.

(scenes 1-7), a near field monitor Genelec 8020c was chosen, because of its flat on-axis frequency response and its relatively small size. For scenes 8-11, in addition to measurements using the Genelec 8020c, an omnidirectional dodecahedron measurement speaker was used for the measurements, with the intention of processing EDCs and calculate room parameters from the measured RIRs. For the purpose of acquiring BRIRs with a sufficiently high signal-to-noise ratio for auralizations and perceptual experiments, PA loudspeakers of type QSC K8 were additionally used for the reference measurements in scenes 9-11. Each sound source is described by its position, its orientation and its radiation directivity. The depicted reference points do not exactly correspond to the (frequency dependent) acoustic center of the sound sources, but were chosen according to the manufacturers information (QSC K8) and the dimensions of the enclosure (Genelec 8020c).

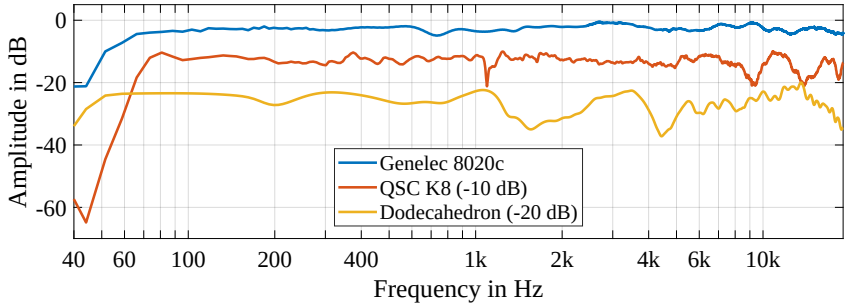


Figure 3.26: On axis frequency responses of the three sound sources. QSC K8 speaker results are shifted by -10 dB, the dodecahedron results by -20 dB for better visibility.

Near-field monitor: Genelec 8020c

The Genelec 8020c (see Fig. 3.25, left) is a commercially available near-field monitor including a 3/4" (19 mm) tweeter and a 4" (105 mm) bass driver. The lower and the upper cut-off frequency, provided in the manufacturer's datasheet, are 65 Hz and 21 kHz, respectively, the maximum peak level is given as 105 dB. This compact active loudspeaker was chosen because it can be easily and accurately positioned and produces sufficient output levels for single-channel IR measurements of the scenes. As its on-axis frequency response (cf. Fig. 3.26) only shows minor variations, it is a suitable device for obtaining high signal-to-noise levels in the corresponding frequency range. Its reference position is defined 18 cm below

the top of the bass driver on the center axis (cf. Fig. 3.25). The loudspeaker is used in all eleven scenes for single-channel IR measurements and in scenes 1, 3 and 5 also for the situations involving a binaural receiver.

PA-Loudspeaker: QSC K8

The QSC K8 (see Fig. 3.25, center) is a commercially available portable powered two-way speaker consisting of a 8" (203 mm) bass driver and a 1.75" (44.5 mm) high-frequency unit. The manufacturer's datasheet documents the frequency response (−6 dB) from 66 Hz to 18 kHz. The measured on-axis frequency response is given in Fig. 3.26. In comparison to the Genelec 8020c speaker, the frequency response shows more variations and despite using a bass driver with a larger diameter, the QSC K8 speaker has a steeper low-frequency slope at a higher frequency than the Genelec 8020c. The QSC K8 was selected in favor of the Genelec 8020c speaker for the binaural measurements of scenes 9, 10 and 11 as it provides higher sound pressure levels (maximum peak level of 127 dB according to the manufacturer).

Measurement speaker: ITA's dodecahedron

The dodecahedron speaker (see Fig. 3.25, right) is a custom DSP-driven three-way system using a single low-frequency driver operating up to 177 Hz, and mid/high-frequency units each consisting of 12 speakers in spherical enclosures (cross-over at 1.42 kHz). The cross-over frequencies correspond to the upper cut-off frequencies of the 125 Hz and 1 kHz octave bands. More technical details are presented in [20].
The dodecahedron with its omnidirectional radiation pattern is used in scenes 8-11 to capture RIRs with the main intention to process room acoustic parameters according to ISO-3382-1 [76]. As the loudspeaker is applied in the three-way setup, meaning that the high frequency unit placed on top of the mid frequency unit which is mounted on the subwoofer, the on-axis frequency response (cf. Fig. 3.26) shows more variations than the other two sound sources. For frequencies between 1 kHz and 5 kHz, this is mainly caused by reflections occurring on the neighboring elements of the speaker.

Directivity measurements and database

As sound sources differ in their spatial radiation, directivity measurements of all three sound sources have been conducted in a hemi anechoic chamber for an equi angular grid of $2^\circ \times 2^\circ$. The dodecahedron and the Genelec 8020c were positioned on a turntable, varying the azimuth angle, at a height of 2 m. The elevation

angle of the measurement position was varied (northern hemisphere only) by an arm that was equipped with a half inch free-field microphone at a distance of 2 m from the sound source. The sound sources were flipped and the measurement was repeated to capture the southern hemisphere. For the dodecahedron, separate directivity sets were acquired for the mid-frequency and high-frequency unit, in both cases including the full physical setup of the three-way system, each referenced to the center of the corresponding unit (cf. 3.26). The low-frequency unit was modeled omnidirectional. The directivity measurement of the QSC K8 speaker was conducted using the ELF loudspeaker measurement system (Four Audio). During this measurement, the microphone was placed on the floor of the hemi anechoic chamber at a distance of 8 m.

Post-processing involved the removal of the propagation delay, a subsonic high-pass filter and time windowing. To increase the spatial resolution to $1° \times 1°$ of the final dataset, a spherical harmonics interpolation [135] was applied separately, to the magnitude and the unwrapped phase spectra. More details on the measurements, post-processing and the final database can be found in the documentation of the BRAS [12].

Directivity patterns of the final directivity datasets in the horizontal plane for four one-third octave frequencies are shown in Fig. 3.27. Increasing directionality for higher frequencies in case of the Genelec 8020c and the QSC K8 speaker is observed, while the dodecahedron maintains its omnidirectionality also for the 3.15 kHz frequency band, despite showing some irregularities up to 6 dB.

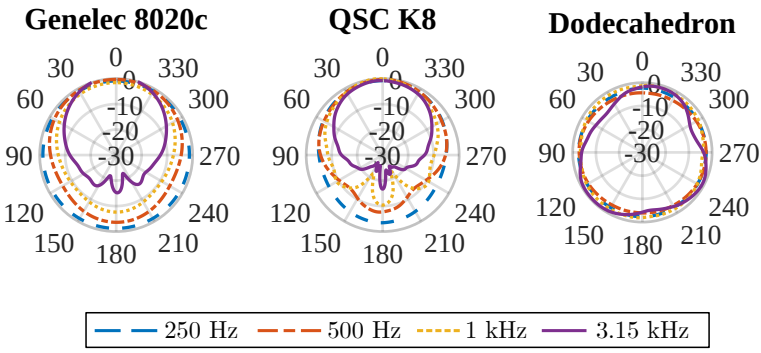


Figure 3.27: Directivity patterns of the horizontal plane for four one-third octave bands of all sound sources

3.5 Receivers

Receivers of the BRAS database are either defined as being omnidirectional or as binaural. While omnidirectionality can be assumed in the scene definition, no actual corresponding measurement device is able to fulfill this condition. For the measurement of single channel IRs of the simple scenes (1-7), a G.R.A.S. 40AF 1/2" free field measurement microphone was used. The manufacturer reports frequency ranges of 5 Hz – 10 Hz and 3.15 Hz – 20 kHz for a variation of ±1 dB and ±2 dB, respectively [59]. While this on-axis frequency response is appropriate for capturing broadband IRs, the variations for other angles of incidence exceed 4 dB at 10 kHz, hence, omnidirectionality is not given. Despite directing the microphones towards the sound sources in the simple scenes, this represents a limitation of the BRAS database as all of the simple scenes contain reflected sound waves coming from an varying angles of incidence.
Because of the reverberant sound field, in the complex room scenes (8-11) a 1/2" B&K 4134 diffuse-field measurement microphone was used to capture the single-channel RIRs, the manufacturer lists its frequency range (±2 dB) from 4 Hz to 20 kHz [42]. Deviations from omnidirectionality of about 4 dB also occur for 10 kHz.
For the binaural receivers in scenes 1, 5 , 8, 9, 10 and 11, the FABIAN head and torso simulator [93, 41, 40], equipped with DPA 4060 microphones, was used. This system allows the measurement of BRIRs for different head-above-torso orientations (HATOs) which later can be applied in listening experiments based on dynamic binaural synthesis. Examples of HRTF pairs in the frequency domain for three directions are presented in Fig. 3.28.

3.6 Geometry

In scenes 1-7, the objects such as reflector panels and wooden profiles were manufactured and positioned with a high level of accuracy. The size of the individual objects showed variations of below ±2 mm, an accurate positioning inside the anechoic chambers was achieved using a laser distance meter (Bosch DLE 50 Professional, precision ±1.5 mm), a self leveling cross line lasers (Bosch Quigo, precision ±0.8 mm/m) and an angle measurement device (geo-FENNEL EL 823, equipped with a Hama LP-21 laser pointer, precision ±0.5°). While in the scene description of the database, all coordinates and dimensions are provided as milimeter values, only a positional accuracy in the range of $1 - 3$ cm be guaranteed, as the challenging conditions of the wire-woven floor in the fully anechoic chamber (scenes 2-4) or matching the reference points of the objects for longer distances of multiple meters did not allow a more accurate positioning in the

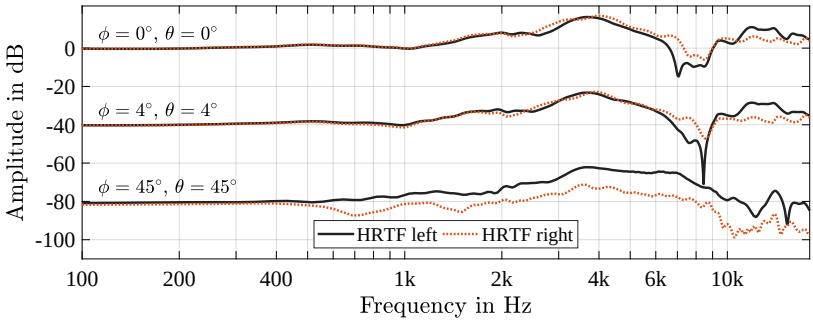


Figure 3.28: Head-related transfer functions (HRTFs) of the FABIAN dummy head system for three directions. Data corresponds to the default head-above-torso orientation (0°).

reference measurements. The acquisition of the geometry for the complex rooms was based on different techniques. For scenes 8, 9 and 10, the initial geometry acquisition was performed with a TOPCON EM-30 laser distance meter (precision ±3 mm) mounted on a *VariSphear* scanning microphone array [24], which allows to position the laser distance meter in azimuth and elevation using two computer controllable motors. This scanning process leads to point clouds, which were then post-processed in a 3D modeling software. Major room surfaces were modeled as planar surfaces by fitting a plane through the corresponding points. As this procedure would have required a very tedious process of scanning a very high number of points in case of scene 11, the geometric data of scene 11 was derived from architectural drawings and was then validated against a set of 15 manually measured distances in the room. In general, the geometry modeling processes aimed at a detail threshold of 50 cm, which is typical for GA simulations [177]. However, for the final models, some details of smaller dimensions were manually modeled (based on manual measurements in the rooms). Examples for such details for scene 9 and scene 10 are given in Fig. 3.29. In case of scene 9, the 3D model (c) contains the lights of the room, but no cable ducts. The stairs to the stage in scene 10 were modeled (d) without the handrail and the metal beams supporting the stairs. simplifications and details which were not accounted for in the final 3D models. In the final modeling step, material surfaces were assigned to the polygons. The database contains a collection of detail photos for each of the room scenes accompanied by a list of simplifications and images that show the omitted elements. Handrails and cable ducts, for example, with diameters of a few centimeters or chandeliers were not modeled.

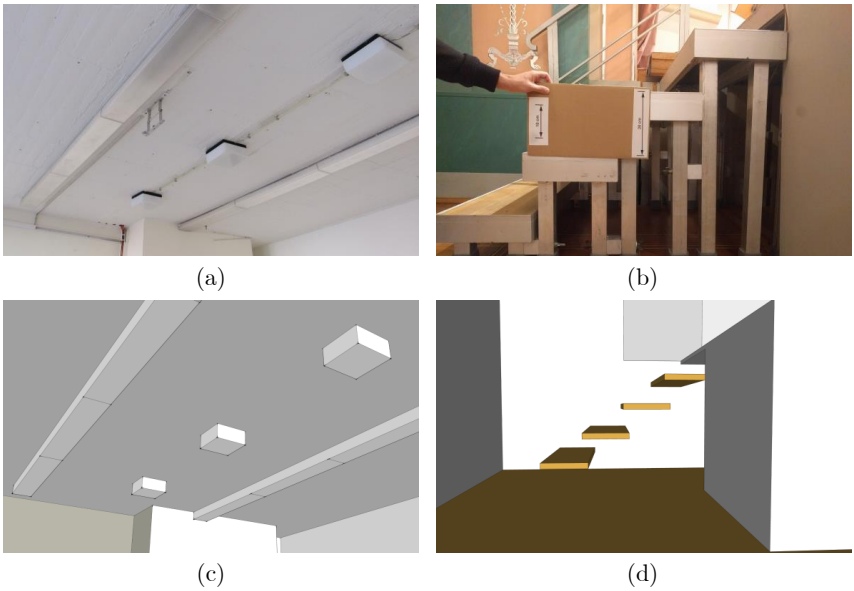


Figure 3.29: Examples of geometry details (a: ceiling of scene 9, b: stairs to stage in scene 10) and their representation in the 3D models (c, d)

3.7 Boundary conditions

In an ideal reference database, all boundary conditions should be described by their angle dependent acoustic impedance $\underline{Z}(\phi, \theta)$. However, neither can all different surfaces with their varying types of installation be measured in the laboratory, nor are any (standardized) full-range measurement techniques available to determine them in situ. For this reason, it was chosen to model the boundary conditions of the BRAS by absorption and scattering coefficients, based on different acquisition methods. The database contains 28 surface materials in total, stored in separate 37 data files including absorption and scattering coefficients in third octaves from 20 Hz to 20 kHz. For some materials, multiple files are available in order to provide data for different angles of incidence.

3.7.1 Absorption coefficients

Simple scenes

Table 3.6 gives an overview of the valid frequency range and the scenes the materials were used in. For scenes 1-7, absorption coefficients are either based on ISO 10534-2 [73] normal incidence impedance tube measurement, or on measurements using the setup of scene 1. The latter method was applied to process angle-dependent absorption data of the stone wool absorber (scene 1 and 2) and the diffuser (scene 1). The lower frequency range of this method (300 Hz and 500 Hz) depends on the height of the reflecting object and thus differs for the two materials, while the higher frequency limit of 15 kHz is caused by the uncertainty with respect to positioning of source, receiver, and the material probe. The ISO 10534-2 measurements are valid above 100 Hz due to mechanical limitations of the loudspeaker, and below 4 kHz due to the diameter of the tube. Details on both measurement and the post-processing can be found in the documentation of the BRAS [13].

Table 3.6: Overview of surface materials, the corresponding absorption coefficient acquisition method and scenes in which materials were used

Material	**Absorption coefficient acquisition method (valid freq. range)**	**Scenes**
Medium density fiberboard	ISO 10534-2, normal incidence (100 Hz to 4 kHz)	1–7
Stone wool absorber	ISO 10534-2, normal incidence (100 Hz to 4 kHz)	1–2
Stone wool absorber	Angle dependent in situ measurement (300 Hz to 15 kHz)	1–2
Wooden diffuser	Angle dependent in situ measurement (500 Hz to 15 kHz)	1
Tiles	Estimated random incidence values	1, 5–7
Room surfaces (24 materials)	Estimated random incidence values	8-11

Complex rooms

For the complex rooms (scenes 8-11), several surfaces were acoustically investigated with a hand-held in situ device [107], which consists of a loudspeaker in a spherical enclosure and a combined sensor unit measuring sound pressure and particle velocity [168]. This method delivers valid results for normal incidence if applied for porous absorbers in controlled scenarios, but faces several challenges if applied on different materials in complex rooms, leading to high uncertainties for reflective surfaces and repeated measurements as well as small and complex objects. Thus, the absorption data was derived from in-situ measurements whenever possible and attributed from material databases otherwise [177, 128]. These datasets are also referenced by the term *initial estimates*.
The database also contains a second absorption dataset of *fitted estimates* for all materials of the complex rooms. For each third octave band, the reverberation times calculated by the Eyring equation T_{Eyring} is matched to the measured reverberation times T_{meas} by applying a frequency dependent correction factor $k(f)$ to all *initial* coefficients of the materials of the corresponding scene (cf. Eq. (3.7)).

$$T_{\mathrm{meas}}(f) = T_{\mathrm{Eyring}}(f) = 0.161 \, \frac{V}{-S \cdot ln(1 - k(f) \cdot \bar{\alpha}') + 4mV} \,, \tag{3.7}$$

with V being the room volume and S the surface area of the scene's 3D model; m denotes the air attenuation [88]. The average absorption coefficient of the initial estimates $\bar{\alpha}'$ was calculated using

$$\bar{\alpha'} = \frac{1}{S} \sum_i S_i \alpha'_i \tag{3.8}$$

where the surface area $S = \sum_i S_i$ occupied by each material was taken from the 3D room models. Solving Eq. (3.7) for k leads to:

$$k(f) = \frac{1}{\alpha'} \cdot \left(1 - e^{(0.161 \cdot V / T_{\mathrm{meas}}(f) - 4mV)/S}\right) . \tag{3.9}$$

Rounding all absorption coefficients to three decimal places and limiting the value range to [0..1] leads to, when re-calculating the reverberation time using the Eyring equation, deviations averaged across all one-third octave bands of less than 0.1% from the measured T20 values. Only for room R_2 of scene 8, a slightly higher average error across the third octave bands of 0.28 % was determined.
The resulting correction factors for the complex room scenes are shown in Fig. 3.30. This data shows that the initial absorption coefficients were chosen as too small for lower frequencies. Correction factors greater than 1 are observed for frequencies around 100 Hz in case of all rooms except for R_2 of scene 8. In the mid-frequency range between 500 Hz to 1 kHz, the correction factor is close to 1

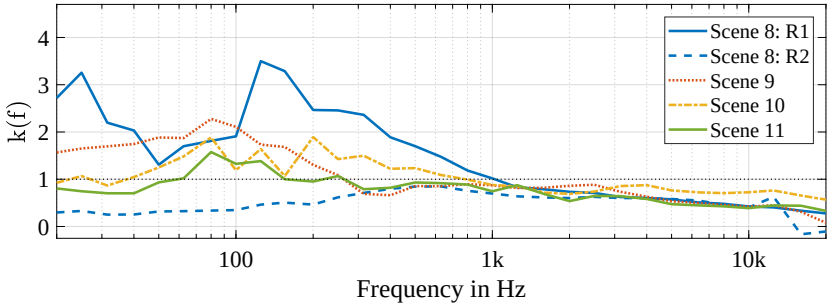


Figure 3.30: Correction factors for absorption coefficients of all complex room scenes. For scene 8, the fitting was conducted for the individual rooms R_1 and R_2. For a correction factor of 1 (dotted horizontal line), the *intial* absorption coefficients remain unchanged.

for most rooms, while in the higher frequency range (above 2 kHz) the correction factor is consistently below 1.
The consistency of the correction factors across the rooms indicate a general mistake during the definition procedure of the intial absorption coefficients. While for the higher frequencies this is surprising, for lower frequencies the underestimation of absorption occurs as the simplified room model leads to less absorption in contrast to the real room environment containing more geometrical structure and small cavities.

3.7.2 Scattering coefficients

To describe the surface structure of the surface materials of the database, random-incidence scattering coefficients are provided for most materials of the database, which are typically accounted for by GA-based simulation software [46]. Despite ISO 17497-1 [74] and ISO 17497-2 [75] include standardized measurements of random-incidence scattering and directional diffusion, no measurements have been conducted. The difficulty to remove material samples from the rooms is one reason, another reason is an increased level of uncertainty when measuring very smooth surfaces such as the MDF plates.
In the first versions of the database (see Table 3.1 at the beginning of this chapter), the scattering data was taken from tabulated values of comparable materials [177] or estimated according to the structural dimensions of the materials [52, 115]. In the most current version of the database [13], the scattering coefficients are calculated according to an estimation function [132] based on the characteristic

depth d_{char} [177] of the material in order to provide a more transparent calculation model for the materials. The estimation function is given by Eq. (3.10).

$$s(f) = 0.5 \cdot \sqrt{\frac{d_{char}}{c/f}} \tag{3.10}$$

with c being the speed of sound and f the frequency. As the scene geometries also contain larger flat surfaces, the lower value range was chosen as 0.05, the upper limit for the scattering coefficient was set to 0.99. The structural depth of all materials is listed in the material description files located in the surface description folder of the database [13]. Scattering data for the 31 one-third octave center frequencies between 20 Hz and 20 kHz are calculated for all materials, except for the wooden diffuser (scene 1), as for this scenario, the modelling of scattering effects of the diffuser was the subject of the investigation.

3.8 Discussion

In this section, the BRAS database is discussed with respect to validity, accuracy and uncertainty of the input data and the reference measurements. In general, it has to be stated that despite enormous scientific and technological advances of humanity in the past few centuries, it is still a great challenge to describe the physical reality with the required detail for an acoustical situation in a controlled laboratory environment. Even in the macroscopic world, the accuracy of positioning loudspeakers, microphones and reflecting elements is limited not only by the accuracy of the measurement device, e.g., a laser distance meter, but also by the researcher who has to reliably use the measurement device and properly document the result. And similar to quantum physics, where the measurement of certain effects is challenging as the measurement itself interferes with the physical effect, the components of the acoustical scenes of the BRAS also have an impact on the sound field of the scenes, i.e., the stands of the microphone or the enclosure of the loudspeakers. While such effects can be neglected for lower frequencies, this is certainly not the case for the highest audible frequencies, which also need to be considered if the database should be applied for perceptual investigations involving audio material covering the full audible frequency range. Thus, it cannot be claimed that the BRAS database contains valid input data and reference measurements for acoustical scenes for the full audible bandwidth from 20 Hz to 20 kHz. While several issues could be addressed, such as a more detailed scene description including physical dimensions and material properties of all elements of the scene, there is a lack of adequate methods to acoustically characterize these elements for the entire frequency range.

Nevertheless, the database represents a useful collection of data for the validation of simulations in academia and in professional practice – if the user is cautious when analyzing and comparing data, and considers the limitations of the database. For simplified measures such as a clarity value of a room in the mid-frequency range, measured results and provided input data of the BRAS database can be trusted, while a detailed spectral or temporal analysis in low or very high frequency bands should not be conducted without discussing the increased level of uncertainty. This is supported by investigations concerning the general uncertainty of room models and validations [178], which, for instance, demonstrate that a change of the room temperature by 1° C leads to completely different fine structure of the room transfer function, while an evaluation in one-third octave bands barely shows any differences.

3.8.1 Input data

Scene geometry and Boundary conditions

For the simple scenes, an accurate description of the scene geometry does not impose many challenges. The combination of uncertainties during the manufacture and the positioning of multiple elements in case of some scenes lead to variations of around ±1 cm, which barely impact the results of acoustical simulation or measurements. The acoustical characterization of the materials, however, involves many problematic issues, such as the selection of an adequate measurement method. The characterization of the acoustical impedance of a material is only standardized for normal incidence using an impedance tube (ISO 10534-2 [73]) and for random incidence based on reverberation chamber measurements (ISO 354 [77]). To create an ideal input dataset for benchmark scenes, angle dependent impedance data is required. While several measurement techniques exists [104, 112, 106] they often lack robustness and flexibility with respect to the measurement setup, are only valid for certain materials types or reflection models and also come with various uncertainties. Fig. 3.31 shows results based on different measurement techniques and demonstrates the level of variation in the absorption coefficient [14]. The lack of inter-laboratory reproducibility of standardized ISO 354 and ISO 10534-2 measurements have been shown in various investigations [63, 173, 72, 131]. The result based on the PU probe (*In situ*) [168] also deviates from results based on the impedance tube measurement according to ISO 10534-2. Similar results of this comparison were obtained in more detailed investigations by Hirosawa et al. [68] and by Pedrero et al. [121]. For the complex scenes, it has to be kept in mind that the definition of the input dataset for the four scenes represents a modeling process. The definition of the

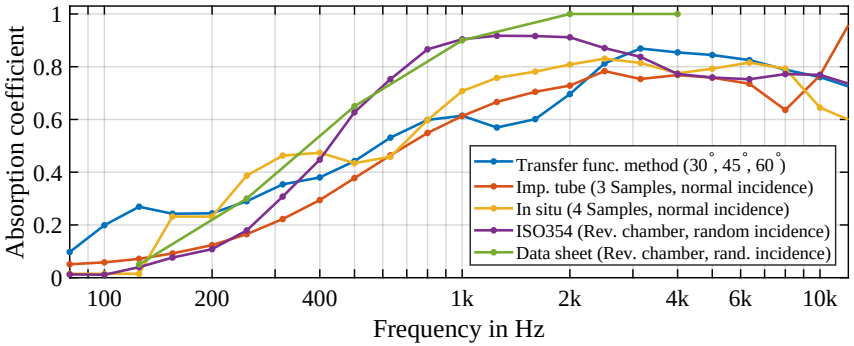


Figure 3.31: Comparison of absorption coefficients of the stone wool absorber (RockFon Sonar G) based on different acquisition methods

room geometry, albeit based on accurate measurement technology, is a subjective simplification process far from being a deterministic process. The same applies for the selection of the selection and the assignment of materials to the surfaces of the generated geometry. Even if an appropriate broad band method for measuring (angle dependent) complex impedances was available, the required number of acoustic measurements would be impracticably high in complex environments. A simplified representation of the acoustic environment using a reduced number of surface materials, and estimated or simplified boundary conditions is thus currently inevitably and remains a subjective process. Cox and D'Antonio describe this as "*[...] not an entirely satisfactory situation, as subjectivity should not be part of a prediction model*" [46]. Avoiding this subjectivity and making the definition of the input dataset a deterministic (and automated) process is especially challenging for GA-based simulations, as the input data of the scenes usually involves a third input parameter, such as scattering or diffusion coefficients and all three parameters, geometry, absorption data and scattering data, are not independent from each other [9]. If the defined geometry of a scene, for instance, includes highly detailed structures of a room, a lower scattering coefficient to the corresponding surfaces would have to be defined – or vice versa. This is, in addition to the challenging conditions when trying to measure elements of a complex room in a laboratory environment, another reason why no measured scattered data (e.g., according to ISO17497-1 [74]) was included in the BRAS database. When considering simulations of higher frequencies using wave-based models, another promising and potentially deterministic approach to define the room geometry would be automated laser scans or photogrammetry [97], which would, however, require several to-be-defined steps of a processing chain from

setting up the acquisition device until the final 3D model data is obtained. As such a technique is based on optical measurements, resulting models would only include room geometry which is visible from the position, the measurements are taken. It would thus neglect coupled volumes connected to the main volume of the room (cf. scene 10).

Sound sources and receiver

The selection of sound sources and receivers for the acoustical scenes also leads to limitations and uncertainties when analyzing and discussing corresponding simulation and measurement results of the scenes.
The missing provided data of the *omnidirectional* microphones and the uncertainty of directivity measurements of the three sound sources and the binaural receiver limits the accuracy of the input data description. The omnidirectionality assumption is less critical for the half-inch microphone capsules, where the first zero in the directivity appears well outside the audible range at an angle of 90°, and a frequency of 27 kHz where the wave length equals the capsule's diameter. To provide more accurate results, the actual directional behavior of the microphones in the higher frequency range could be studied by repeating measurements of different microphone orientations.
Another source of error is the positioning of the sound sources and the receivers. Differences between the geometrical and acoustically estimated positions were up to 5.5 cm, which corresponds to approximately 7 samples at a sampling rate of 44.1 kHz, and a frequency of 6.2 kHz. While this is uncritical with respect to changes of the overall level and also has a negligible influence on the measured room acoustical parameters [159, 187], this uncertainty has to be kept in mind when comparing temporal and spectral information of simulated and measured data.

3.8.2 Frequency range limitations

While the database offers input data for the full audible frequency range from 20 Hz to 20 kHz, the validity of this input data and of the reference measurements of the scenes is restricted by a substantially smaller frequency range. Fig. 3.32 depicts the frequency ranges for the boundary conditions, for the description and definition of the sound sources and receivers, and the reference measurements included in the database. The presented frequency ranges are not strict limits, but thresholds based on different criteria. The solid lines indicate validity of the provided data, a dashed line is used if this validity is slightly reduced and the dotted line represents severely reduced validity of the data. For the boundary conditions, the lower and the higher limit of the applied ISO 10534-2 method

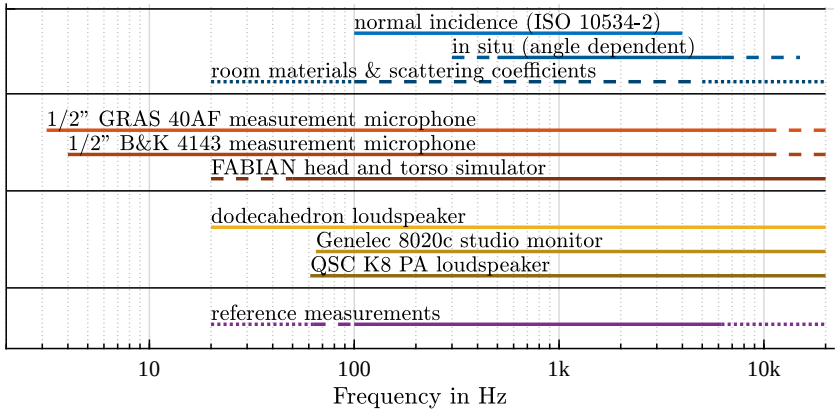


Figure 3.32: Frequency limitations of boundary conditions, receivers, sound sources and the reference measurements. Dashed lines indicate that the data might have slightly compromised validity, dotted lines indicate severely reduced validity.

corresponds to the lower frequency cut-off of the loudspeaker and impedance tube diameter, respectively. The in situ measurement method is restricted by edge diffraction effects, which are relevant up to 300 Hz–500 Hz, and the reduced accuracy when positioning loudspeakers and microphones, which affects the absorption data for frequencies between 6.2 kHz and 15 kHz. The absorption data of the room materials and the scattering data of all materials is not directly based on measurements and thus is fully described by dashed and dotted lines. According to the general availability of absorption data in databases[2] and literature [46, 177], the frequency range of slightly compromised validity is defined as 100 Hz to 5 kHz [148], which also roughly corresponds to the frequency range of the standardized absorption measurement techniques (ISO 354 and ISO 10534-2). In contrast to the boundary conditions, the receivers and their description in the database have a wide frequency range and impact results only for high frequencies above 10 kHz, where omnidirectionality can not be assumed anymore (variations of up to ±4 dB). The FABIAN dummy head including its high spatial resolution is only limited by its microphones (20 Hz) and the measurements (50 Hz, [38]). The sound sources, which are also characterized by the high resolution directivity data, limit the analysis only by their lower reproduction limit.

[2] 98% of the material entries (2369 in total) in the PTB absorption coefficient database [128] include data at least from 125 Hz to 4 kHz

The limitations of the receivers and the sound sources have to be accounted for when analyzing the reference measurements, which are additionally impacted by the measurement environment. For the simple scenes, the lower cut-off frequency of the measurement rooms is 63 Hz and 100 Hz for the full anechoic chamber (scenes 2, 3 and 4) and the hemi anechoic chamber (scenes 1, 5, 6 and 7), respectively. If a detailed temporal or spatial analysis involving the reference measurements is conducted, this analysis should consider frequencies below 6.2 kHz due to the positioning uncertainty of receivers and sound sources in all scenes, but particularly in scenes situated in the fully anechoic measurements (2, 3 and 4), where the positioning of sound sources and receivers was most challenging.
In summary, this section reveals substantial constraints, mostly with respect to the boundary conditions, when working with the database, which originally aimed at providing a reference dataset valid for the full audible frequency range. Nevertheless, simulated data based on the provided input data can still be analyzed and compared to the reference measurements, even out of the valid frequency ranges, as long as higher uncertainties of input data and reference measurements are cautiously considered during the comparison. With respect to the room materials, it is the author's hypothesis, that is impossible to define valid absorption and scattering coefficients for a variety of simulations software whenever these tools use different simplifying approaches.

3.8.3 Reproducibility of measurements

In general, the applied transfer function measurement method using exponential sweeps [103] leads to reproducible measurement results, as long as all elements of the measurement chain are linear and time-invariant, i.e. describe a LTI system. While linearity can be assumed for the applied equipment and the generated output levels, there is a potentially relevant amount of time variance in the measurement chain, especially in the room itself, caused by variations of temperature, humidity, and in case of the complex rooms, furniture and loose objects which do not have permanently fixed positions. Even though the temperature and the humidity of a room can be measured and documented during the measurements, it is not possible to exactly recreate these conditions within the physical environments of the BRAS scenes. Without a deeper analysis, this section presents examples comparing results of repeated measurements.

Typically room acoustic parameters such as the reverberation time are rather insensitive to the aforementioned slight violations of the LTI system. The chamber music hall of the *Konzerthaus Berlin* (scene 10) was also part of a scientific investigation comparing chamber music halls in 2004, conducted by Hidaka and

Nishihara [67]. The documented value of the EDT (1.32 s, average for 500 Hz and 1 kHz octave bands) in the unoccupied hall exactly matches the corresponding EDT evaluated based on the BRAS measurements. The averaged clarity value C_{80}, however, depends on the chosen positions of sound sources and receivers[3] during the measurement and shows a deviation of more than 0.5 dB. Hidaka and Nishihara report in their overview a value of 2.0 dB (averaged for 500 Hz, 1 kHz and 2 kHz), while the BRAS measurements lead to an averaged value of 1.5 dB.

With respect to the evaluated reverberation times, the dodecahedron measurements of the BRAS database can also be compared to measurements taken with the Genelec studio monitor in the complex room scenarios. Despite of the studio monitor becoming directional for the mid-range frequencies, the energy decay of the room should be rather independent from the sound source. The averaged reverberation times (500 Hz and 1 kHz) of the dodecahedron speaker and the Genelec speaker for scene 10 only deviate by less than one percent (T_{20}: 0.16%, T_{30}: 0.48%). Measurements using the Genelec speaker were conducted on the same day, but around two hours later than the dodecahedron measurements of scene 10.

For the RIR measurements of scene 9, the exact setup of the BRAS database was tested three days earlier. The results of the test measurements compared to the measurements included in BRAS database for one sound source and one receiver position in the frequency domain are visualized in Fig. 3.33. The magnitude

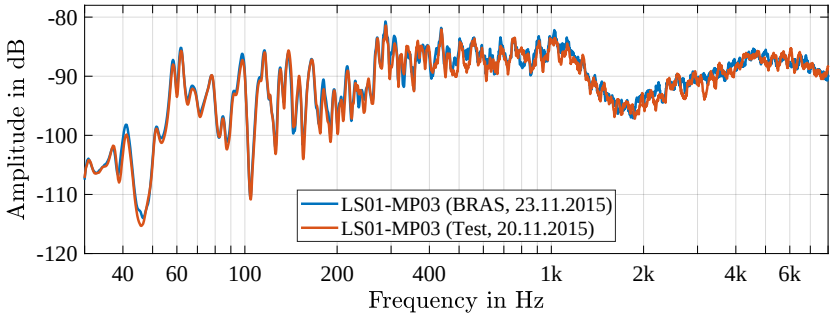


Figure 3.33: Room transfer function of test measurements and BRAS measurements of scene 9, taken on different days. A smoothing filter (1/24 octave moving average) was applied to the result.

[3] Unfortunately the exact measurement positions are not documented in [67]

frequency responses of the RIRs only deviate marginally from each other, the deviations slightly increase for higher frequencies. This is mainly caused by the limitation in reproducing the exact positions of sound sources and receivers. An analysis of all other sound source and receiver positions lead to similar results. These limitations related to the reproducibility should be considered when comparing simulated results to the measurements.

Without providing further details, more examples of deviations observed for repeated measurements along with an estimation of the general measurement uncertainty are listed in Table A.1 (Appendix A.2). The combined general measurement uncertainty in the time domain for RIRs is calculated according to Witew [188] based on the *Guide to the expression of uncertainty in measurement* (GUM) [78].

4

Round robin comparison for uninformed simulations

Parts of this chapter have been published in: Brinkmann, F., Aspöck, L., Ackermann, D., Lepa, S., Vorländer, M., and Weinzierl, S. (2019). A round robin on room acoustical simulation and auralization. The Journal of the Acoustical Society of America, 145(4), 2746-2760 [37].

In the tradition of the previous three round robin comparisons of room acoustical simulation [176, 29, 30, 31], a new round robin competition has been organized in which the participants were not informed about the measured results of the reference scenes. For this reason, this investigation is called either *uninformed* or *blind* comparison. In contrast to the first three round robins, which each discussed only one space, the investigation presented in this chapter includes in total nine different acoustical environments, which are a part of the database discussed in Chapter 3. In addition to room acoustic parameter analysis of the complex rooms, this investigation also features a temporal and spectral analysis of (room) impulse responses the participants had to provide for each of the nine scenes. As participants also submitted binaural room impulse responses for some scenes, listening experiments based on binaural auralizations could be conducted. A related study including listening tests is summarized in Section 4.5 and presented in more detail in [36].

4.1 Method

To evaluate results obtained by different room simulation tools, the organization team[1] announced a call for the round robin on room simulation and auralization in March 2016. This call was directed to users and developers of room simulation

[1] This team consisted of researchers from TU Berlin and RWTH Aachen University and was a part of the research unit *Simulation and Evaluation of Acoustical Environments* (SEACEN) funded by the DFG (German Research Foundation)

tools of both, academia and industry and distributed, and advertised in related mailing-lists, at international conferences and on the project website [158]. The simulation input dataset containing a general documentation, scene descriptions and models of nine different scenes, receiver and sound source characteristics, and boundary conditions, was provided to the participants of the round robin. This dataset corresponds to version 1 of the database presented in Chapter 3.
In contrast to the round robins I–III, where participants had to provide room parameters for one or more octave band center frequencies, in this investigation the participants were asked to submit single-channel and binaural impulse responses, in order to increase the evaluation possibilities and avoid variations caused by different parameter processing [29, 81]. Thus, a spectral and temporal analysis of all results including the comparison to the measured results is possible. In the analysis in this chapter, absolute magnitude and arrival times of the impulse responses are neglected. Such systematic deviations are taken out by means of normalization and temporal shifting of participant's impulse responses. Applied post-processing methods are explained for each scene separately in the corresponding subsection.
This chapter analyzes simulated and measured impulse responses (and related processed parameters) of the nine scenes with respect to the following research questions:

1. How much do simulated results obtained by different users, who were not informed about the measured results, and different software differ from the measured results?

2. How much do the simulated results differ from each other?

Related to this, the discussion of the results also involves explanations or potential reasons for the discovered deviations. As simulation software typically is a complex environment allowing for various simulation configurations (e.g., ray tracing particles) and input options of the acoustically relevant data (e.g., data format of source directivity), the user has to be considered as a part of the simulation system, also including potential mistakes made by the human user when preparing the simulation.

4.2 Participating simulation tools

After the call for the round robin on room simulation and auralization had been announced, in total 22 groups were interested to participate in this organized comparison. However, only in total seven participants were able to submit results before the deadline in February 2017. One of these seven participants retracted

his submission after discovering errors in the simulation. The six remaining teams and their applied different simulation algorithms are listed below [37]:

- BRASS [170] is a ray tracing algorithm developed in academic research, which clusters reflections up to fifth order to provide accurate early reflections without deploying an image source model.
  Submission by: Developing team, University of Rio de Janeiro
- EASE (Version 4.4) [1] is a commercial tool for the simulation of room acoustical and electro-acoustical environments, using image sources for the direct sound and early reflections, and ray tracing for the late reverberation.
  Submission by: Experienced user, Technical University Berlin
- ODEON Combined (Version 14.02) [115] is a commercial tool for room acoustical simulation based on a combination of the image-source method with a modified ray tracing algorithm.
  Submission by: Developing team, Odeon A/S
- ODEON Combined (Version 12.0) [115] – see above.
  Submission by: Experienced user, University of Bologna
- RAVEN (Version: 2017a) [151] uses a hybrid algorithm combining image sources for the direct sound and early reflections, and a ray tracing using the diffuse rain technique for the late reverberation.
  Submission by: Experienced user, RWTH Aachen University
- RAZR (Version 0.90) [185] is an open source software for the simulation of rectangular rooms through a combination of image sources, and a feedback delay network for late reverberation.
  Submission by: Developer, University of Oldenburg

All algorithms are based on geometrical acoustics and consider frequency dependent absorption and scattering coefficients, air absorption, and exchangeable receiver and source directivities – with the exception of RAZR, which assumes omni-directional sources, and does not account for scattering (at the time of submission). Diffraction simulation is only implemented in ODEON [140], and is only applied in case of a blocked direct sound path. The simulations were either conducted by the software developers themselves (BRASS, ODEON v14 and RAZR), in the other cases by experienced users. To avoid a bias, the RAVEN simulations were conducted by a person who was not aware of the measurement results. The results of these teams are presented anonymously in this chapter, labeled with the letters *A*-*F*.
Developers of wave-based algorithms reported that a simulation was either infeasible as source and receiver directivity could not be easily accounted for, or too

time-consuming for the entire audible bandwidth, especially for scenes involving complex geometries. One research group published results for three scenes obtained with a BEM simulation, at least up to mid-range frequencies between 1 kHz and 2 kHz [65]. The team of researchers who retracted their results later published their corrected data independent from the round robin evaluation [161, 166]. These results, however, are not accounted for in this chapter, as they were not officially submitted and, by the time of publication, the authors were already informed about the measurement results.

4.3 Results of the simple scenes

In this section, simulated and measured results of the simple reference scenes in the time- and in the frequency domain are presented in various figures. In case of scenes containing numerous sound source – receiver combinations, only a selection of interesting results are shown for the sake of brevity. As some participants did not model correct arrival times[2] and absolute levels, the direct sound of each impulse response was shifted to the theoretical arrival time according to the distances and the speed of sound. This improves the possibility to visually compare the results in the time domain. A level normalization was applied by dividing each simulated impulse response by its *rms* value. With respect to time windowing and application of a high-pass filters, the post-processing of the simulated results was conducted as described in Section 3.2.

Scene 1 (RS1): Single reflection (infinite plate)

For scene 1, five teams submitted results which are included in the evaluation. Simulated results of these five contributions compared to the measured result in the time domain are shown in Fig. 4.1. For the rigid floor reflection (top), all simulated impulse responses include two distinguishable impulses, the direct sound at 12.3 ms, corresponding to the distance of source to receiver (4.23 m) followed by the floor reflections arriving at 17.5 ms and 17.0 ms, for the combination LS2–MP2 and LS1–MP3, respectively. In case of *F*, the reflected impulse occurs around 0.5 ms earlier, which indicates a mistake of the user during setup of the simulation, e.g., a wrong position of receiver and/or loudspeaker. While the distance between direct sound and reflected impulse is correct for most of the participants, the temporal structure of the direct sound and the floor reflection, which is primarily impacted by the impulse response and the directivity of the loudspeaker, only coincides with the measured impulse response in case of

[2] The results of three participants consistently contained temporal delays of up to 100 ms in all submitted impulse responses

participant E and to some extent with participant A. The measured reflected

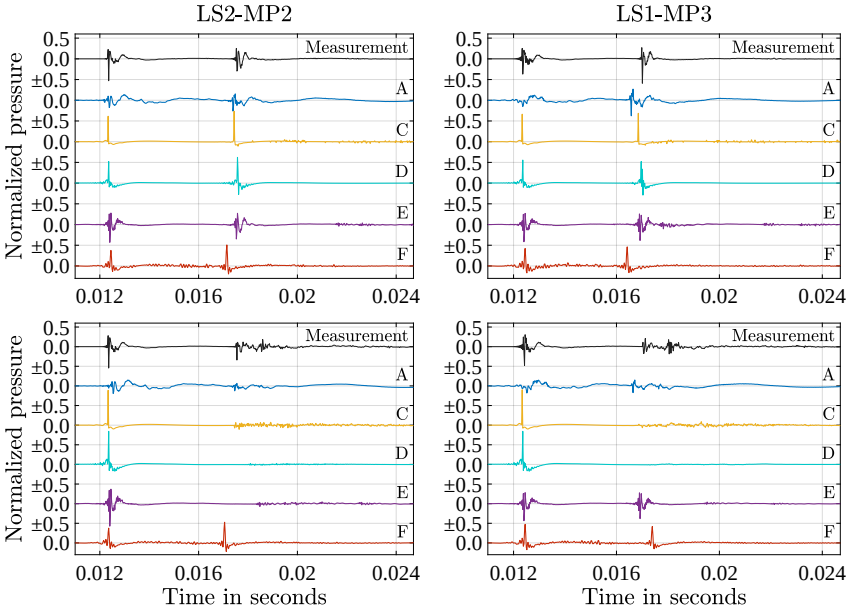


Figure 4.1: Measured and simulated impulse responses of scene 1, for two sound source – receiver combinations and two different reflecting surfaces (top: rigid floor, bottom: diffuser)

impulse of the wooden diffuser (Fig. 4.1, bottom) is less distinct and the energy is spread over a longer period in time. For LS2–MP2 this is also the case for all simulated data except for F, but none of the simulated impulse responses resemble the measured curve. This is similar for LS1–MP3, except for the data of E, which shows a more distinct reflection arriving earlier than the scattered measured reflection. This indicates that for this source–receiver combination, an audible image source on the wooden diffuser could be found, while for LS2–MP2 this image source was absent.

The corresponding results in the frequency domain are presented in Fig. 4.2. For the rigid reflection (top) between sound source LS2 and receiver MP2, a comb filter structure is clearly visible and a good agreement between measurement and simulation is observed for C and E and, to some extent, also for D. The results of A exhibit a low-pass characteristic with a higher magnitude in the lower frequency range than in the mid-frequencies. The temporal inaccuracy of F leads

to a shift of the peaks and notches of the comb filter. Similar results as for the rigid reflection are observed also for the reflection on the absorber, the corresponding plot is included in the Appendix in Fig. B.1. Here, a good agreement for the analyzed frequency range is only achieved by *C* and *E*. The measured transfer functions in case of the diffuser (cf. Fig. 4.2, middle and bottom) show a more irregular comb filter, which could not be matched by any of the simulations. For LS1–MP3, results of *E* and *F* only slightly resemble the measured result.

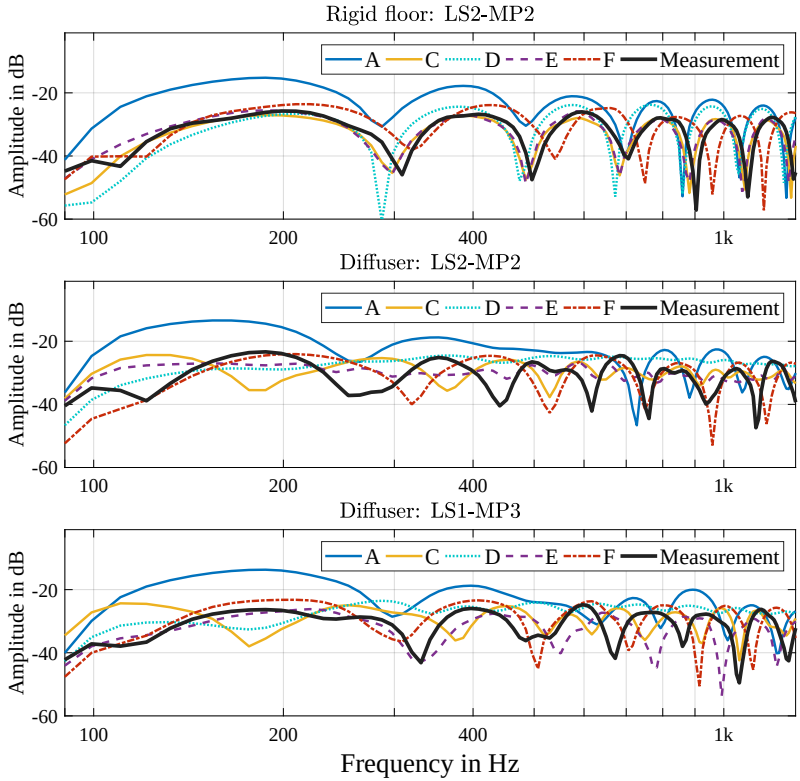


Figure 4.2: Measured and simulated transfer functions of scene 1 for three different situations

Scene 2 (RS2): Single reflection & diffraction (finite plate)

For scene 2, it is particular interesting in how far the simulation algorithms are able to account for the finite size of the reflectors. Four participants submitted results for this scene, the time domain data for the reflection on an absorber panel (1 m x 1 m) is shown in Fig. 4.3. The graph on the left contains the

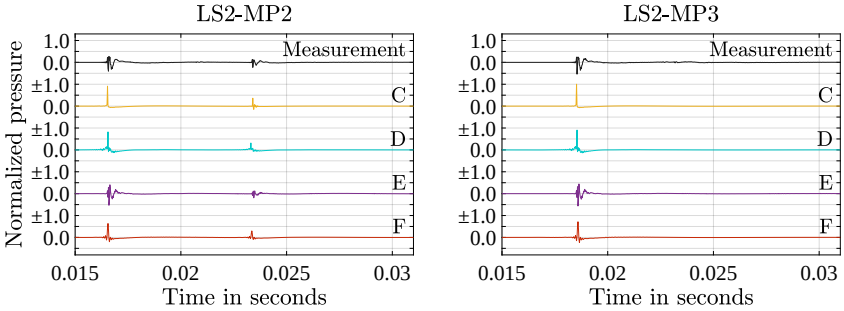


Figure 4.3: Measured and simulated impulse responses of scene 2: reflection on absorber (1 m x 1 m) for two sound source – receiver combinations

impulse responses for LS2–MP2, corresponding to a specular reflection with $\phi_{in} = \phi_{out} = 45°$. Despite the absorption of the panel, in case of all impulse responses of this configuration, a reflection at 23.4 ms is visible, mostly containing low frequent energy (cf. Fig. 3.31 for the absorption coefficient of the absorber). For the measured result and for all simulation results this reflected impulse has a smaller amplitude than the direct sound, which was not the case for the corresponding reflection on the rigid floor of scene 1 (cf. graph at top left in Fig. 4.1). Thus, the simulation software are able to model the reduced reflected energy, however, an analysis in the frequency domain is required to check whether the level of the reflection is accurately modeled. In case of LS2–MP3, the reflection is vanished for all simulated results, the measurement shows a very weak impulse caused by diffracted and scattered sound waves. The impact of this weak reflection can be seen more clearly in the frequency domain. The corresponding graph, see plot in the middle of Fig. 4.4, contains a weak comb filter effect in the frequency range from 200 Hz to 1500 Hz, which none of the simulation programs are able to correctly model. In case of the specular reflection on the absorber panel, the results in the top graph show that C is able to account for the limited size of the panel, as the comb filter effect vanishes for lower frequencies. However, the results of C still contain significant deviations from the measured frequency

response. The results of D, E and F do not account for the limited size of the panel at all and model a strong specular reflection which therefore leads to a more pronounced comb filter structure, also in the lower frequency range. The same behavior for these three participants is observed for the reflection on the rigid panel (cf. bottom Fig. 4.4), only C accounts for the effects introduced by the limited size of the panel for frequencies below 1 kHz and shows less deviations from the measurement than in the case of the absorbing panel.

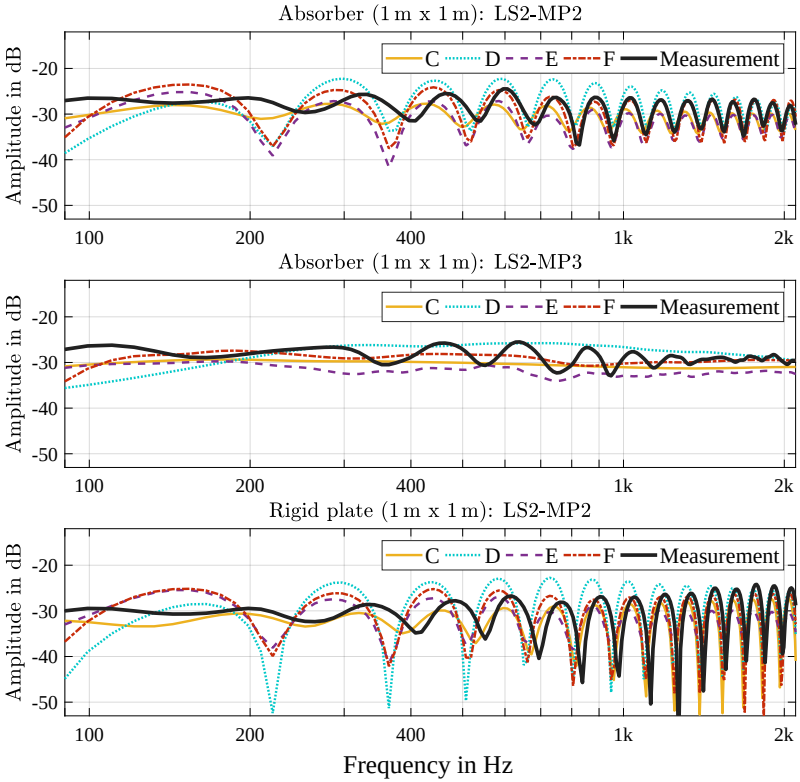


Figure 4.4: Measured and simulated transfer functions of scene 2 for three situations

Scene 3 (RS3): Multiple reflection (parallel finite plates)

In scene 3, only one sound source and one receiver position is defined, located between two parallel plates (2 m x 2 m) which are positioned 10 m apart from each other. Due to the outcomes of scene 2, a significant error of the simulations for low frequencies can be predicted as all algorithms except for *C* do not account for the limited size of reflecting panels. The repetition of these reflections on finite sized panels therefore lead to larger errors for higher reflection orders. As the scene's sound source, the *Genelec8020c* speaker, is directed on one of the panels, correct directivity modeling is also crucial for the simulation of this scene.
The measured and the simulated results of the five participants, who contributed data in this case, are presented in Fig. 4.5. In Section 3.3 it was demonstrated

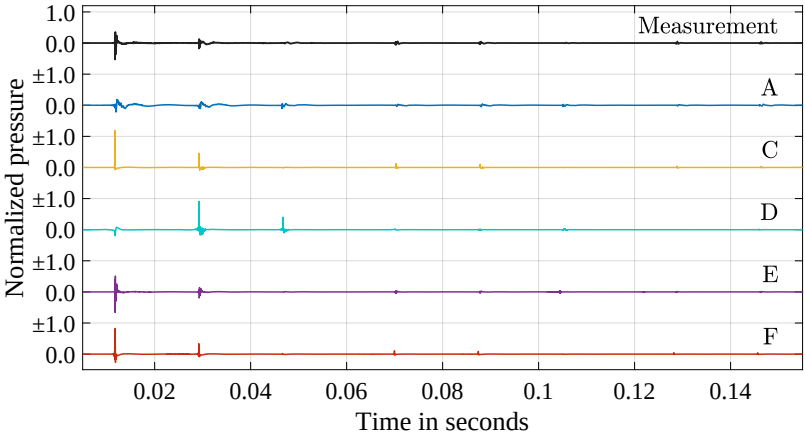


Figure 4.5: Measured and simulated impulse responses of scene 3, LS1–MP1

that it is straightforward to apply an image source model to temporally match the reoccurring reflections of the measured impulse response. This is achieved by all contributing participants, whereas the levels of the reflections are not correctly modeled by any of the participants. The result of *D* includes a first order reflection with a higher amplitude than the direct sound. This is caused by a wrong orientation of the sound source in the simulation[3]. In the initially provided documentation of the scene, the organization team accidentally included ambiguous scene information about the sound source orientation, which was

[3] In the simulation of *D*, the sound source was directed facing the wall and not facing the receiver

later corrected. Participant D, however, missed to correct the simulation results accordingly. In the impulse response graphs with the linear y-axis shown in Fig. 4.5, A, E and F appear to have acceptable results with respect to the amplitude of the reflections. Plotting the results with a dB scale, however, reveals substantial errors when comparing to the measured reflections. 4.6 shows two examples of the less deviating results, the data of participant C and of E, both compared to the measurement. Despite the normalization, the direct sound of

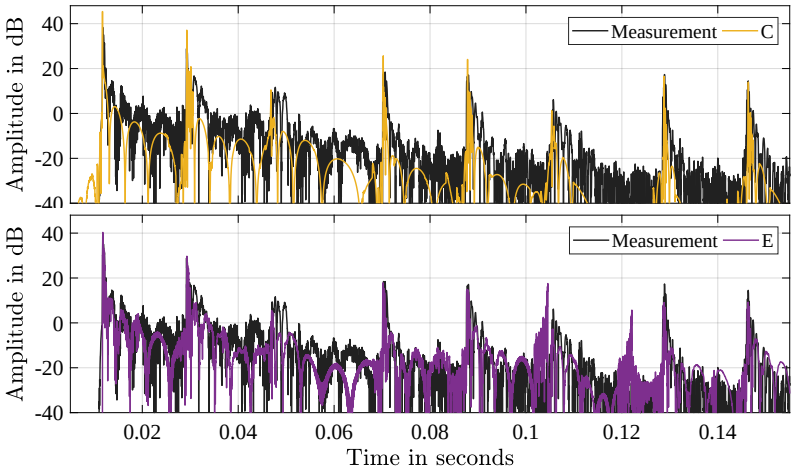


Figure 4.6: Measured and simulated impulse responses (logarithmic y-axis) of scene 3, LS1–MP1. Simulated results only for participants C and E.

participant C has a higher amplitude than the measurement. This deviation is caused by the *rms* normalization – the measurement contains more energy by weak reflections in the anechoic measurement room, either by the room or by small objects such as microphone or loudspeaker stands. Furthermore participant C did not apply the impulse responses of the Genelec8020c loudspeaker, but the only the provided magnitude spectrum, potentially without the provided phase information. This leads to symmetric impulses with a linear phase, which can be visualized by only plotting the direct sound without application of the high-pass filter (see Appendix B, Fig. B.4). This explains the peak-level deviation of the simulated impulse response from the measurement in general, nevertheless, when analyzing all impulses of the first 160 ms, the peak level differs inconsistently from the measured peaks, partly being smaller, partly being larger than the peaks of the measured impulse response. For participant E, the simulated impulse

responses resembles the measured curve to some extent, especially for reflection 1, 3 and 4. This is an indication, that the impulse responses of the loudspeakers and not just energy-based magnitude spectra were applied in the simulation. However, there is a substantial deviation for reflection 2 and the simulated impulse response contains an additional reflection at 120 ms, which is neither present in the measurement, nor in the result of C and the simple image source model (cf. Section 3.3, Fig. 3.5).
As this impulse response is mostly characterized by the reoccurring reflection on the MDF plate the loudspeaker is directed towards, it is therefore worthwhile to analyze this first reflection (at 30 ms) more closely. When evaluating the peak level difference between direct sound and the first reflection, the theory of a ideal point source and a pure reflection on a rigid panel of infinite size leads to a drop of 8 dB, neglecting the reflection on the back wall, which contains substantially less energy due to the directivity of the speaker in the measurement situation. In the measurement the peak-level difference is 9.5 dB, the deviations from the theoretic value are caused by the finite size of the reflector and the (low) absorption of the panel, and to a marginal extent due to air absorption for very high frequencies. In the two presented simulation results, the peak level between direct sound and first reflection drops by 8 dB and 10 dB for C and E, respectively. This represents an acceptable error, but as further reflections depend on this first order reflection, the error propagation leads to increased deviations for higher order reflections.
The results in the frequency domain (see Appendix B, Fig. B.3), show that while both simulated results resemble the measured curve to some extent for higher frequencies above 500 Hz, in the lower frequency range the measured results contains a more regular comb filter structure than the simulated data. This deviation is especially prominent for E, which does not account for the finite size of the reflector and thus contains relevant energy in the lower frequency range also for the higher order reflections, leading to a irregular spectrum also in the low frequencies.

Scene 4 (RS4): Single reflection (reflector array)

Scene 4 was designed in a way that if purely GA-based simulations only calculate specular reflections, the first configuration (on-center reflection) corresponds to a single reflection on a small rigid panel (cf. scene 2, for results see Fig. B.2), while the second configuration (off-center reflection) would not deliver a reflection at all. This was indeed the case for one of the three participants who submitted results for this scene, and potentially a reason for other participants not to submit results at all. The measured and simulated results for two situations, on- and off-center reflections for $\phi_{in} = \phi_{out} = 45°$ are shown in Fig. 4.7. The measured impulse

response for the on-center situation consists of the direct sound and a reflection impulse with the reflection having a higher peak level than the direct sound. For the three simulation results, this is the case for D only and the temporal structure of the impulses does not resemble the measured curve for any of the simulations. However, the results of scene 2 revealed that D did not account for the finite size of the reflectors, and thus, in case of scene 4, only processed the reflection panel in the center, treating it as a panel of infinite size and neglecting the acoustical effect of all other array elements. For the off-center case this applied model of D is even more problematic. Due to the reflection point being in between the plates, the impulse response only contains the direct sound, while the measurement as well as A and C contain a slightly different reflection with respect to the on-center situation. Therefore A and C apply models which account for reflection objects in the vicinity of the specular reflection point. This data plotted in

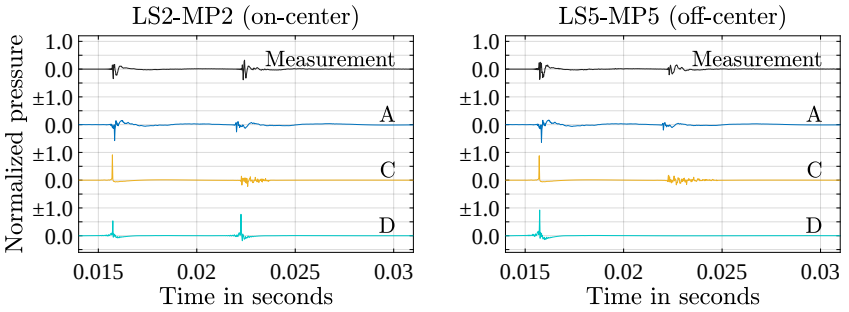


Figure 4.7: Measured and simulated impulse responses of scene 4. Reflection point located in center of middle panel (LS2–MP2, on-center, $\phi = 45°$) and between panels (LS5–MP5, off-center, $\phi = 45°$).

the frequency domain, as shown in Fig. 4.8, allows a more accurate analysis of the reflection models used by the three simulations. The transfer functions of the on-center situation reveals that no algorithm matches the varying notch depth of the measurement. Similar results are achieved for frequencies above $f_g(\phi = 45°) = 1993\,\text{Hz}$, for lower frequencies the reflection in case of C vanishes, while for A and D the comb filter is mostly frequency independent, even for very low frequencies. For frequencies lower than $f_{g,total}(\phi = 45°) = 173\,\text{Hz}$, the energy of the measured reflection considerably dropped and, in this frequency range, matches the curve of C. The data of the off-center situation, depicted in the lower graph of Fig. 4.8, confirms that the impulse response of D does not contain a reflection at all. The transfer function of A and C resemble the measured transfer

function to some extent, but fail to accurately model the reflection especially between 700 Hz and 2 kHz, where the reflection contains a substantial amount of energy (cf. Section 3.3, Fig. 3.9) despite the specular reflection point being between the elements of the reflector array.

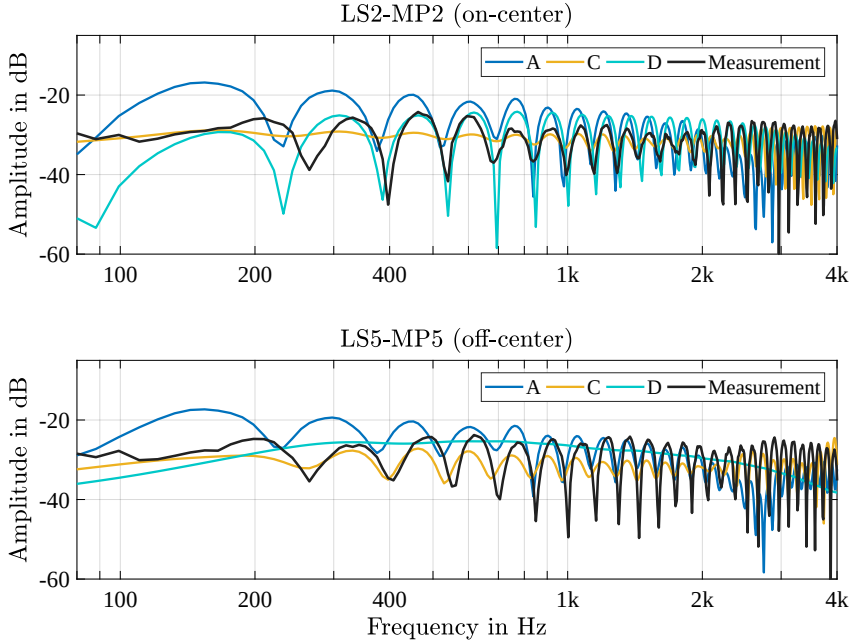


Figure 4.8: Measured and simulated transfer functions of scene 4 for the on- and off-center configuration ($\phi_{in} = \phi_{out} = 45°$ for both situations)

Scene 5 (RS5): Diffraction (infinite wedge)

In total three participants submitted results for scene 5, however, one participant did not account for diffraction effects at all and is therefore excluded from the evaluation. Because the introduction of the participants described which of the contributions were able to model diffraction (cf. Section 4.2), in this Section, the two simulation results are presented anonymously with a different labeling to avoid that the participants can easily be identified.
The data of two situations in the frequency domain are presented in Fig. 4.9. For both situations, there is no line of sight connecting the sound source and the

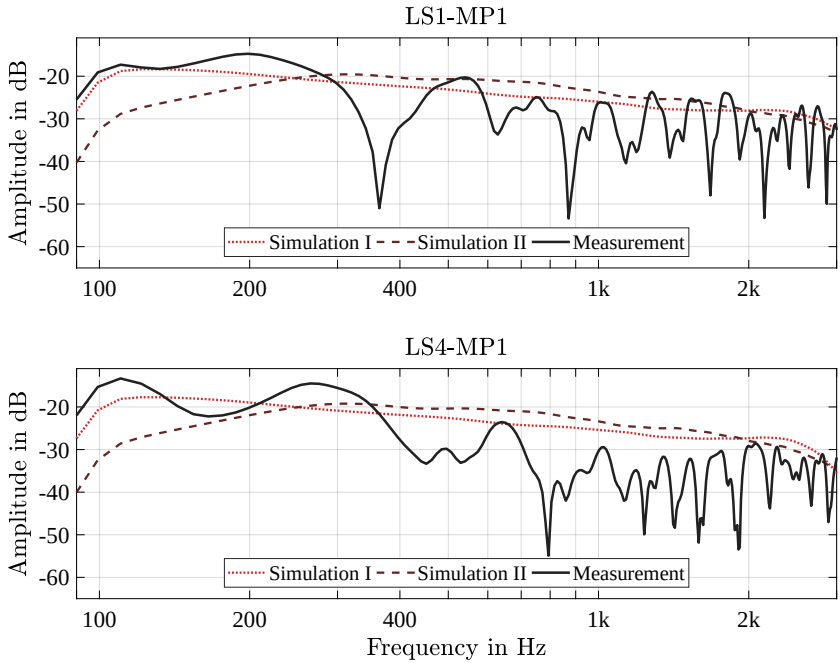


Figure 4.9: Measured and simulated impulse responses of scene 5 for two sound source positions and receiver position MP1

receiver. The receiver is located at position MP1 at a height of 1.24 m behind the partition, which has a height of 2.07 m. The sound source LS1 is located also at a height of 1.24 m, creating a symmetrical situation, while LS4 is positioned on the ground of the hemi anechoic environment. While the measurement shows the impact and relevance of diffracted reflection paths, the simulation only account for first-order diffraction, which typically leads to a frequency dependent attenuation with a low-pass characteristic of the direct sound. The results displayed in the time domain (cf. Appendix B, Fig. B.5) confirm that only the diffracted direct path is accounted for by both simulations. Interestingly, the simulated results hardly differ when comparing the two sound source positions. This is different for the measurements, where additional diffraction paths lead to more variations in spectrum and the impulse response.

4.4 Results of the complex scenes

Scene 8 (CR1): Coupled rooms (laboratory room & reverberation chamber)[4]

For this scene, six participants submitted single-channel RIRs, all based on simulations applying the *initial* absorption and scattering coefficients (see Section 3.7.1). The simulated and measured EDCs are processed using the *ITA-Toolbox* in two steps : First, a bandpass filter for the 1 kHz octave band is applied to the RIRs using the *ita_mpb_filter* method. In the second step, the EDC is calculated using the *ita_roomacoustics_edc* method.
The resulting EDCs of both door angles for the 1 kHz octave band are presented in Fig. 4.10. No participant was able to simulate data resulting in curve resembling

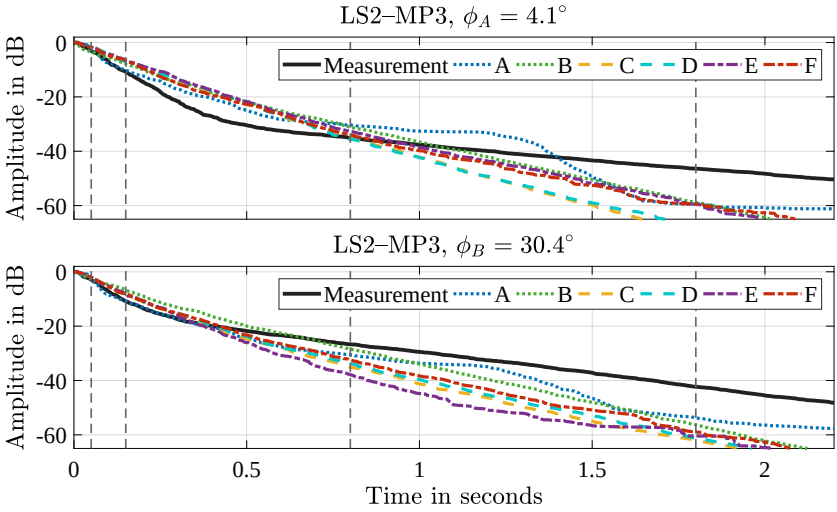


Figure 4.10: Measured and simulated energy decay curves for LS2–MP3, evaluated for the 1 kHz octave band. Dashed vertical lines indicate the early and the late time frame for the decay time evaluation.

the measured EDC for the provided boundary conditions. Visually, only the curves of participant *B* and to some extent also of *A* show a double slope while

[4] Parts of this section have been published in: Aspöck, L., and Vorländer, M. (2019). Simulation of a coupled room scenario based on geometrical acoustics simulation models. In Proceedings of Meetings on Acoustics (Vol. 36, No. 1, p. 015002). Acoustical Society of America [17].

the other participants simulated RIRs for which the evaluated EDCs resemble a linear decay. The decay times (DT), determined as described in Section 3.3 and listed in Table 4.1, confirm the high deviation of the simulation results from the measured data, but also show consistently higher values for $\mathrm{DT_{late}}$ than for $\mathrm{DT_{early}}$. It can also be observed, that especially for A, but also for B the energy decays irregularly. This should be considered when analyzing the resulting decay times. The average ratios $\mathrm{DT_{late}/DT_{early}}$ of all participants are 1.78 and 2.36 for ϕ_A and ϕ_B, respectively, while the measurements have much higher ratios of 7.25 (for ϕ_A) and 5.20 (for ϕ_B). Noteworthy are also the large differences among the participants, e.g., a factor of 2 between the results of $\mathrm{DT_{early}}$ of A and B. This inconsistency among the participants indicates that the simulation of the given scene is not only sensitive to the boundary conditions, but also strongly depends on the configuration and the applied simulation software.

		Meas.	**A**	**B**	**C**	**D**	**E**	**F**
ϕ_A	$\mathrm{DT_{early}}$ [s]	0.73	0.79	1.49	1.02	1.07	1.30	1.28
	$\mathrm{DT_{late}}$ [s]	5.29	1.72	2.16	1.77	1.88	2.27	2.28
ϕ_B	$\mathrm{DT_{early}}$ [s]	0.74	0.75	1.59	0.96	1.03	1.03	0.94
	$\mathrm{DT_{late}}$ [s]	3.85	2.28	2.14	2.18	2.23	2.90	2.34

Table 4.1: Decay times of measured and simulated data for scene 8

Based on the 3D model and provided absorption coefficients, the reverberation times of the individual rooms $\mathrm{T_{Ey,R2}}$ and $\mathrm{T_{Ey,R1}}$ can be calculated using the Eyring equation, the ratio r is then calculated according to Eq. (4.1).

$$r = \frac{\mathrm{T_{Ey,R2}}}{\mathrm{T_{Ey,R1}}} = \frac{2.49\ \mathrm{s}}{0.79\ \mathrm{s}} = 3.15. \tag{4.1}$$

Thus, the ratio of reverberation times based on the provided data of the round robin is substantially lower than the reverberation time ratio of the individual room measurements $r_{meas} = 7.43$ (see Eq. (3.5)). This indicates that the boundary conditions, which were defined without knowing, or referencing to, measured results, lead to inadequate input data for the simulations explaining the large differences to the measured EDCs.

Scene 9 (CR2): Small room (seminar room)

While the geometry of the seminar room is fairly simple, the validity of the simulation results highly depends on the accuracy of the provided input data. Particular interesting with respect to the simulation results of this scene are the lowest frequency bands, as the the simulations are mostly based on GA models

which have a limited validity below the scene's Schroeder frequency $f_s = 233\,\mathrm{Hz}$[5]. The evaluation of this scene includes a validation of the simulation's direct sound arrival times and an analysis of the simulated reverberation time and clarity values. Additionally, a spectral analysis of the measured and simulated BRIRs is presented.
Note that in all plots (including markers) of scene 9, 10 and 11, the markers of the simulated results have been slightly shifted on the x-axis to improve the visibility.

Arrival time analysis

To check whether the participants correctly entered the source and receiver positions, the time of arrival of the direct sound of all ten simulated RIRs is compared to the theoretical time of arrival according to the distance between source and receiver. The times of arrival of the simulated RIRs were determined using the ITA-Toolbox method *ita_start_IR* and are presented in Fig. 4.11. With the exception of *F* for LS2-MP1 and LS2-MP2, all simulated results show identical or only slightly deviating arrival times. However, for in total three participants systematic time delays in the simulated RIRs were observed: 86 ms for *C*, 12 ms for *D* and, in case of six of the ten RIRs, 83 ms for *F*. These systematic delays are independent from the simulated positions and were corrected, which lead to in general very low deviations from the expected arrival times for nearly all simulations. All deviations are below 10 samples, except for *F*, who did not correctly define the scene when simulating RIRs for LS2-MP1 and LS2-MP2.

Room acoustic parameters

Room acoustic parameters of the simulated RIRs were evaluated for full octave bands ranging from 125 Hz to 4 kHz, using the method *ita_roomacoustics*[6], which is available in the ITA-Toolbox. In Fig. 4.12, the simulated EDT and T20 values are compared to the measured values including the JND of ±5% indicated by the dashed lines. Despite the JND being only an appropriate measure for a subjective difference limen if the results of individually processed RIRs are discussed, the JND is included as a reference also in the analysis of averaged results.
No participant provided RIRs which lead to averaged EDT values between the JND limits for all six octave bands. In case of the three octave bands 500 Hz, 1 kHz and 2 kHz, the results of the participants are more consistent. Especially for 1 kHz, five simulations resulted in similar values slightly below the measured

[5] Calculated using $f_s = 2000\sqrt{T_{\mathrm{avg}}/V}$ for V=146 m^3 and T_{avg} =1.98 s.
[6] For the simulated RIRs, the noise compensation method *noCut* was selected, the measured RIRs were processed using the method *cutWithCorrection*

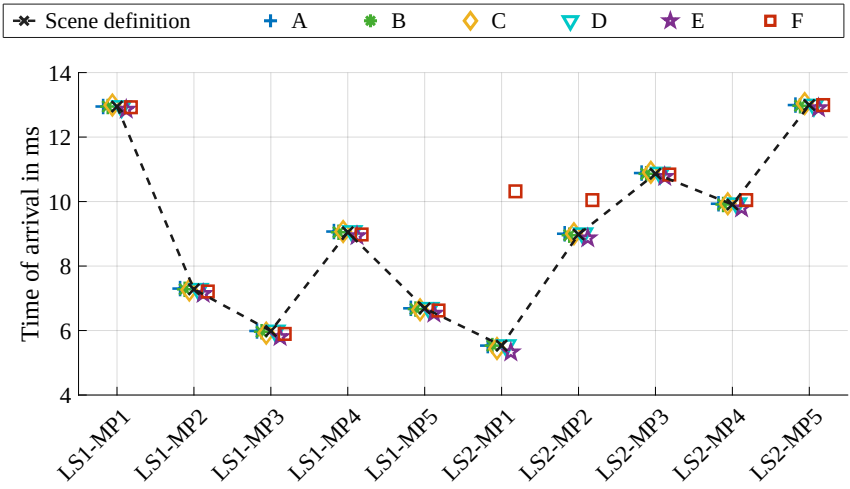


Figure 4.11: Evaluated arrival times of all simulated room impulse responses for scene 9. In case of three participants (C, D and F), a constant time shift was applied as results included systematic delays.

value. The values of A deviate the most from the measured results (absolute mean deviation: 0.71 s) and decrease for higher frequencies, whereas the results of C, E and F are similar in all frequency bands. Participants B and D have the lowest absolute deviation of 0.17 s and 0.21 s, respectively, averaged across all six frequency bands. Interestingly, most simulated results of the 125 Hz octave band substantially deviate towards a higher EDT value, while the values of B and D are relatively close to the measured results for this frequency band. In comparison to the EDT, the simulated reverberation times T20 deviate less from the measured values and include rather consistent results among the participants for 500 Hz, 1 kHz and 2 kHz with deviations from the measurement slightly below or above the JND (with the exception of A for 2 kHz). The results of participant B are consistently close to the measured values for all frequency bands with the lowest absolute deviation (0.13 s averaged across the six octave bands) of all participants. In addition to the simulated results, the graph showing the T20 values also includes the reverberation time predicted using the Eyring equation, indicated by the dotted gray line. These values are calculated using the room volume and the surfaces of the 3D model, and the corresponding absorption coefficients. Apart from the simulated data for 2 kHz and 4 kHz of A, all Eyring based values are lower than the measured and simulated values for frequencies

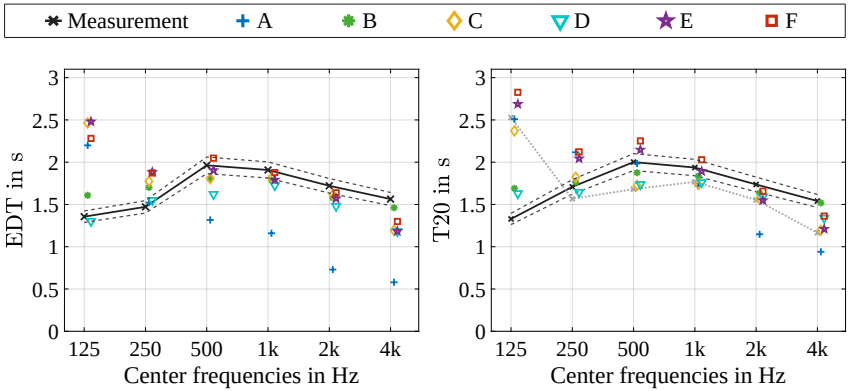


Figure 4.12: Early decay time and reverberation time for scene 9, averaged for ten evaluated room impulse responses. Dashed lines indicate the just-noticeable difference of 5 %, the dotted gray line in the plot on the right represents the reverberation time calculated with the Eyring equation.

between 250 Hz and 4 kHz. The very high value 2.5 s for 125 Hz is an indication that the provided boundary conditions for this room might be erroneous and thus leads to the high deviation of four simulation programs and the equation based value.

Processed clarity values for two evaluated RIRs are shown in Fig. 4.13. The measured curve for LS1-MP3 shows a drop of the clarity values from 4 dB at 125 Hz to −1 dB at 1 kHz, before the value slightly increases above 0 dB for 2 kHz and 4 kHz. Neglecting the 125 Hz octave band, this trend is also visible for all simulation results. For participants *C*, *D*, *E* and *F* the simulated values between 250 Hz and 2 kHz do not exceed the JND of 1 dB with respect to the measured data. Similar to the observations made for EDT and T20, the largest deviations of the simulation programs are also observed in the 125 Hz octave band and highest deviations can be registered for *A*. This holds true also for the other evaluated RIR of sound source LS2 and receiver MP2. For both presented RIRs, all simulation results, except for *A*, consistently exceed the measured value by 2 dB in the 4 kHz frequency band, indicating potential problems with the simulation model, the geometry data or the boundary condition of the scene.

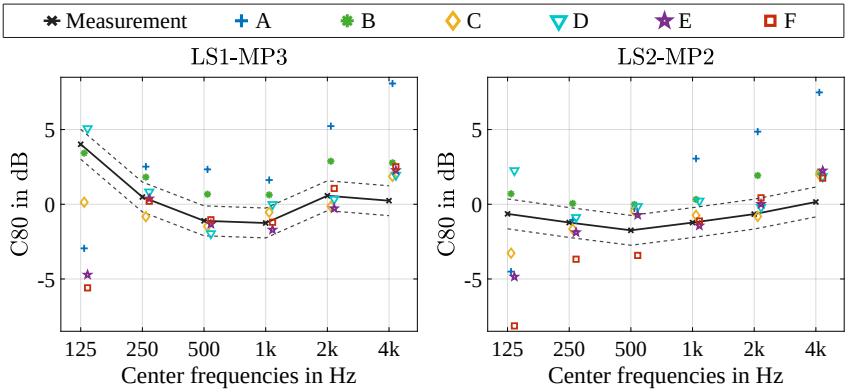


Figure 4.13: Clarity values for scene 9, evaluated for two room impulse responses. Dashed lines indicate the just-noticeable difference of 1 dB.

Spectral analysis

Substantial deviations of reverberation time and clarity values from measured values were observed for all participants, most of all in the lowest and highest evaluated octave bands. These deviations are mainly systematical and occur independent of the investigated sound source and receiver position. The extent of these deviations suggests that differences are clearly audible when comparing auralizations based on simulated and measured RIRs. Without providing a more detailed analysis or an experimental investigation, this statement is supported by the evaluation of the scene's BRIRs in the frequency domain. The spectral difference of the binaural room transfer functions (left channel) for source position LS7 and receiver position MP6 is depicted in Fig. 4.14.

All four simulation results considerably deviate from the measured BRIR. The largest deviations, similar to the outcomes of the parameter analysis, occur for A, and for B deviations of more than 5 dB are observed for frequencies above 1.5 kHz. Closest to the measurement are the results of C and E with spectral differences of around ± 3 dB for frequencies between 500 Hz and 7 kHz. The analysis of the right channel of the receiver, as well as the other four receiver positions does not reveal substantially different results when comparing the simulated spectra to the measured spectrum. The analysis of the early part of the BRIR up to 50 ms, presented in Fig. B.6 in Appendix B, indicates that relevant spectral differences occur in the early part of the BRIR and thus lead to perceivable coloration in auralizations. Informal listening sessions of the simulated and measured BRIRs

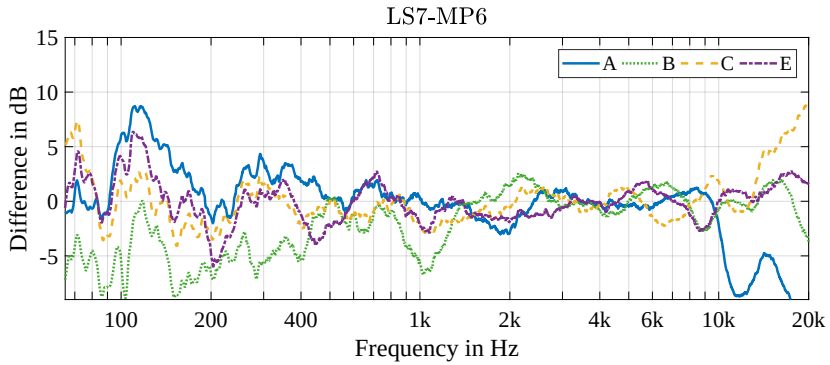


Figure 4.14: Differences of measured and simulated spectrum of binaural room impulse responses (left channel only) for LS7–MP6. A one-third octave smoothing filter was applied to both results before the spectral difference was calculated by means of division in the frequency domain.

convolved with pink noise bursts revealed distinct coloration for all participants. The coloration was particularly strong for participants A and B and lower, but perceivable for C and E. In the direct comparison of C and E with the measurement, the longer reverberation time in the low-frequency range could also be noticed.
A study about differences of the simulated and measured BRIRs involving listening experiments is presented in [37] and summarized in Section 4.5.

Scene 10 (CR3): Medium room (chamber music hall)

In contrast to scene 9, the small concert hall of the *Konzerthaus Berlin* represents a real-world environment, which is used on a regular basis in the condition it is examined in this investigation. The only difference of the room's appearance is the lack of audience and musicians. Nevertheless, the room has a high level of complexity with respect to the surface structure and materials. Additionally, the ceiling and the stage room are also connected to acoustically relevant volumes behind them. These two properties of the room make a simulation more challenging in terms of definition of the input data (boundary conditions and geometry) and the selection of an adequate simulation model.

Arrival time analysis

In order to validate the definition of the sound source and receiver positions the participants entered in their simulation environments, a time of arrival analysis has also been conducted for scene 10, applying the same methods as for the previous scene. The results of this analysis are shown in Appendix B in Fig. B.7. After application of the time shift for participants C and D, the arrival times of the simulated RIRs roughly correspond to the arrival times according to the scene definition. Only for F, the arrival times are either too low in case of six RIRs or considerably higher in case of four RIRs. The arrival times slightly vary from one RIR to another without including a correct relative delay from one position to another. An analysis of the time delay between direct sound and the floor reflection, however, showed that these delays of F match the values of all other participants as well as the delay calculated based on theoretical arrival time of the corresponding image source.

Room acoustic parameters

This section analyzes in how far room acoustic parameters can be accurately predicted in case of a more complex room. In comparison to the previous scene, the greater volume of this scene and the relatively low reverberation time is in general more beneficial for GA simulations with the Schroeder frequency $f_s = 40\,\mathrm{Hz}$[7] being considerably smaller than the lowest evaluated frequency band.
The first parameters presented are EDT and T20, evaluated for all measured and simulated RIRs. The data, averaged across all ten RIRs of the scene, is depicted in Fig. 4.15. Both result plots show that all simulations deviate towards higher values for the lower frequency bands, 125 Hz, 250 Hz and 500 Hz. It is also observed that the range of the simulated EDT values decreases with frequency (for tabulated values, see Table B.1 in Appendix B). With the exception of participant B, the variance of the simulated results decreases further and also deviates only marginally from the measured values for the 1 kHz, 2 kHz and the 4 kHz frequency band. The simulated T20 values vary to a similar extent as the EDT values, but deviate even more from the measured values for the lowest frequency band. For 1 kHz for half of the participants, the deviations are substantially larger than the JND . Only for 2 kHz and 4 kHz, there is less variance among the participants and at least three participants provided data for which the averaged values lie within the JND range of the measured average. When comparing the calculated Eyring reverberation time with the simulated data, it is observed that the Eyring equation leads to lower values for all frequencies, which was also the case for

[7] Calculated using $f_s = 2000\sqrt{T_{\mathrm{avg}}/V}$ for V=3,330 m^3 and T_{avg} =1.31 s.

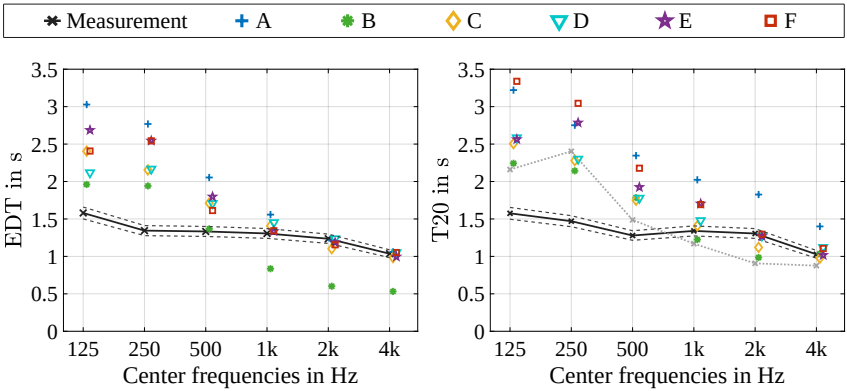


Figure 4.15: Early decay time and T20 for scene 10, averaged for ten evaluated room impulse responses. Dashed lines indicate the just-noticeable difference of 5 %, the dotted gray line (plot on the right) corresponds to the Eyring reverberation time.

scene 9. The substantial overestimation of the Eyring equation for the lowest two frequency bands indicates that the boundary conditions were not correctly defined for the lower frequencies.

In contrast to the reverberation time parameters, which only depend on the decay rate of the energy decay curve, the clarity parameter also depends on the geometry of the room, which directly impacts the arrival time of the reflections in the room and thus determines if a reflection belongs to the early or late part of the RIR . The parameter C80 evaluated for one RIR (LS1–MP1) is presented in the left plot of Fig. 4.16. Most striking is the deviation of B, not providing any value within the JND of ± 1 dB. All other participants show deviations towards lower clarity values for 125 Hz, 250 Hz and 500 Hz, but mostly produced values within the JND range for the three higher frequency bands. A similar trend is observed for the parameter center time, which was chosen due its relevance for concert halls, describing the balance between clarity and reverberance. The simulated and measured center time for LS1–MP1 is depicted in Fig. 4.16 on the right. Here, all simulations except for B resulted in values within the JND range for the frequency bands 1 kHz, 2 kHz and 4 kHz. Similar to what was observed for EDT, T20 and C80, the variance and the deviation of the results is the highest for the 125 Hz and 250 Hz frequency bands.

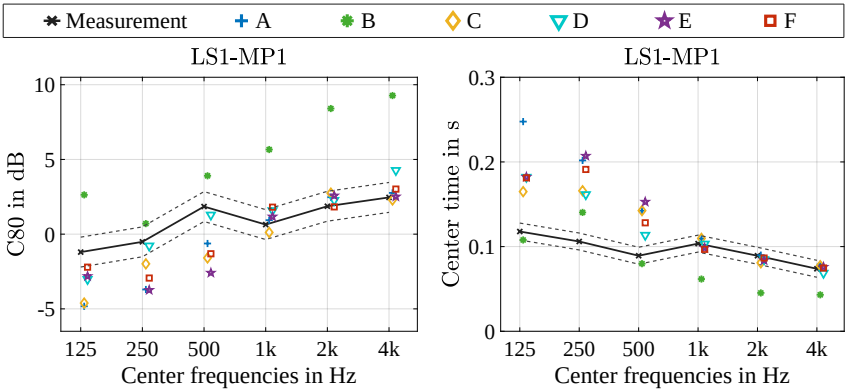


Figure 4.16: C80 (left) and center time (right) evaluated for LS1–MP1 of scene 10. Dashed lines indicate the just-noticeable difference of 1 dB (clarity) and 10 ms (center time).

Scene 11 (CR4): Large room (auditorium)

Because of the rather simple geometry and the volume of the auditorium investigated in this scene, it can be assumed that even software based on simple GA models can achieve acceptable results for this room, as long as correct data of the boundary conditions is provided. Due to the large volume of the room, the Schroeder frequency $f_s = 31$ Hz[8] is the lowest of the three investigated rooms. The evaluation presented in this section includes a temporal analysis and compares simulated and measured EDT, T20 and C80 values.

Temporal analysis

The arrival time analysis was conducted in the same way as for the other two room scenes and revealed similar results as for scene 10, with the exception that in the RIRs of *A* for LS2 the direct sound arrived around 5-15 ms too early in case of all receiver positions (cf. Fig. B.9 in Appendix B). This indicates that possibly the position of sound source LS2 was not correctly entered by participant *A*.
Despite the RIRs of this scene being primarily intended for the evaluation and discussion of room acoustic parameters, an analysis of the early part of the measured and simulated RIRs is additionally presented in this section. In Fig. 4.17, the first 80 ms after the direct sound arrival are plotted for LS1-MP1, for all six

[8] Calculated using $f_s = 2000\sqrt{T_{\text{avg}}/V}$ for V=8,657 m^3 and T_{avg} =2.1 s.

participants and the measurement. It should be kept in mind that all simulations typically include distinct impulses for the direct sound and the reflection as no directivity data for modeling the sound source and the receiver was provided. Since the dodecahedron source used in the measurements does not act like a perfectly omnidirectional point source, the measured impulse response does not show, in contrast to all simulated results, any clearly isolated impulses. However, it can be observed, that a group of eight distinct reflections, arriving between 55 ms and 70 ms, are found in the measured RIR as well as in the simulated RIRs. Only *B* shows a different reflection pattern with substantially less reflections in this time frame. The gap of around 29 ms between the direct sound (and the floor reflection) and these reflection groups is relatively large, as receiver position MP1 is located in the center of the room and LS1 is also located in the front part of the stage with the distances to the back and one of the side walls of around 6.5 m and 8.0 m, respectively. A striking difference between the measurement and the simulation is the temporally spread reflected energy between 30 ms and 55 ms and between 70 ms and 110 ms, which is observed for the measurement, but in case of the simulations, these reflections have either lower levels (*C*, *D*, *E* and *F*) or are not present at all (*A* and *B*).

Room acoustic parameters

The highest reverberation times, compared to the other two room scenes, were measured for scene 11. As shown in Fig. 4.18, the measured and simulated EDT and T20 values are above 2 s for the three lowest frequency bands. Similar to the observations made for scene 9 and scene 10, the highest deviations across all frequency bands are seen for *A* and *B*. When comparing frequency bands, the highest deviations across all participants occur in the 125 Hz frequency band. Acceptable results with low deviations and only slightly varying results among all participants are only observed for 500 Hz and 1 kHz in case of EDT values and only for 1 kHz in case of the T20 values. Different to the results of scene 10, for the 2 kHz and the 4 kHz frequency bands, the simulations consistently return substantially lower values for both parameters. The trend of overestimating the reverberation time in the lowest frequency band and underestimating it for frequencies higher than 1 kHz is also observerd for the Eyring reverberation time, depicted by the gray dotted line in the plot on the right hand side of Fig. 4.18. This indicates inaccurate boundary conditions for the frequency bands with high deviations.

The averaged clarity values of all ten RIRs obtained by the simulation tools compared to the corresponding averaged values of the measurement (see Fig. B.12 in Appendix B) might lead to the impression, that simulations by participants *C*, *D* and *F* produce acceptable results with deviations only slightly greater than the

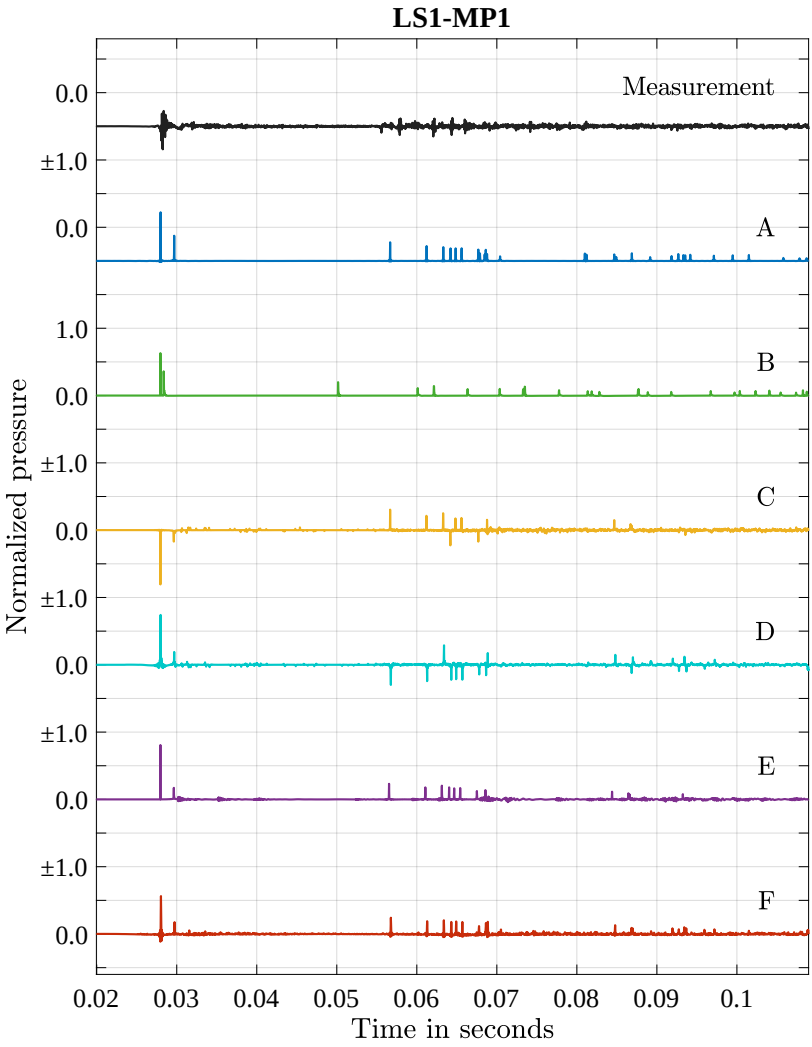


Figure 4.17: Early part of measured and simulated impulse responses for LS1-MP1 of scene 11

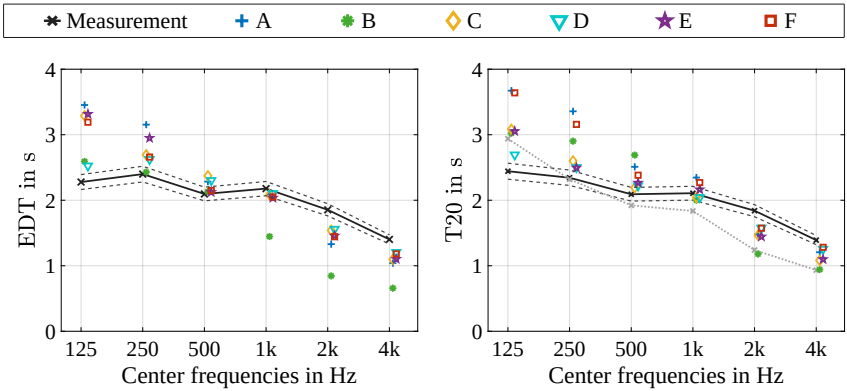


Figure 4.18: Early decay time and reverberation time for scene 11, averaged for ten room impulse responses. Dashed lines indicate the just-noticeable difference of 5 %, the dotted gray line (plot on the right) represents the Eyring reverberation time.

JND for frequencies between 250 Hz and 2 kHz. The clarity parameter, however, includes the direct sound of the impulse response and thus strongly depends on the positions of sound source and receivers. An analysis of the individual RIRs revealed that the deviations of simulation from measurements are not consistent among the different RIRs. This is demonstrated in Fig. 4.19, including the results for LS2–MP2 and LS2–MP3. While for the first position, almost all simulated values lie out of the JND range, the results are substantially closer to the measured data for LS2–MP3. Remarkable is the result of participant *F*, which includes deviations of more than 1 dB for all frequency bands for the first position, but matches the measured values (except for the 2 kHz) for the second investigated position.

4.5 Listening experiments

Based on the BRIR simulations provided by *A*, *B*, *C* and *E* for scenes 9, 10 and 11 (CR2, CR3 and CR4), auralizations were generated and applied in listening experiments using headphone reproduction based on dynamic binaural synthesis [93]. These experiments investigated the *authenticity* [39] and *plausibility* of the simulated BRIRs and, in a more comprehensive experiment, evaluated ten qualities of the Spatial Audio Quality Inventory (SAQI) [94] which are selected according to their relevance for the tested scenarios. The three experiments and

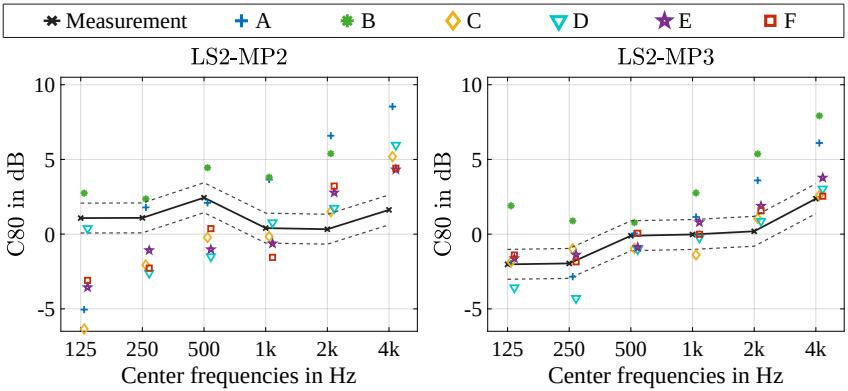


Figure 4.19: Clarity values for scene 11, evaluated for two room impulse responses. Dashed lines indicate the just-noticeable difference of 1 dB.

its evaluation are presented in detail in [37]. This section briefly introduces the experiments and summarizes the results.

Authenticity

In the authenticity experiment groups of seven subjects for each software were asked if they can distinguish measurement and simulation of scene 10 in a two alternative forced choice test (2AFC) using pink noise pulses. The evaluation revealed that all software-based auralizations included clearly audible differences when directly compared to the auralization based on the BRIR measurements. Considering the deviations observed in the parameter analysis, (cf. Section 4.4) the results of this experiment were not surprising.

Plausibility

In the plausibility experiment, in total 28 subjects were assigned to four groups, each corresponding to one software. The subjects had to listen to simulation- and measurement-based auralizations of one sound source position in the medium-sized room (scene 10/CR2) using 20 different anechoic audio stimuli convolved with BRIRs. To measure plausibility, the subjects had to answer the question "*Was this an audio example from a real room?*" after listening to a randomly selected auralization of which the subjects did not know whether it was based on a measurement or simulation.

The evaluation revealed that simulation B was perceived as plausible by all subjects of the group, while for simulations C and E, one of the seven subjects detected artifacts. As suspected when analyzing the deviations of the simulations from the measurement in case of scene 10, the subjects gave the most negative answers in case of the auralization of A, with three subjects perceiving it as implausible. A simulation was considered as implausible if the subjects were able to achieve corresponding 2AFC detection rates of more than 75%, which denotes the perception threshold while 50% denotes the guessing threshold.

Auditory qualities

In this experiment 29 subjects compared all simulation-based auralizations of all three rooms to the corresponding measurement-based auralizations. Ten qualities from the SAQI were selected and explained to the subjects before the test was conducted. The subjects were instructed to compare the simulated data to the measurements using a continuous slider for each of the SAQI qualities with the scale labels as described in [94] displayed above and below each slider. Two audio stimuli, pink noise pulses and a string quartet based on the four off-center source positions in scenes 9, 10 and 11, were selected for the experiment. The auralizations of the string quartet are included in the supplementary material of [37].
The evaluation of the SAQI experiment revealed perceived differences between the different algorithms and between the two audio stimuli. Deviations in 88% of the experimental trials for A are in accordance with the observations made during the spectral and parameter analysis of the room scenes. Softwares B, C, and E performed better and only showed deviations in 12% to 30% of the experimental trials. The lowest deviations from the measurements were observed for E and the string quartet in the small and medium room while substantial deviations for B, C, and E were mainly observed for *difference*, *tone color bright/dark*, and *source position* in case of the pulsed pink noise. In general, the subjects noticed more differences for the pulsed pink noise than for the musical piece. These observations are confirmed by an analysis of variance (ANOVA), which also revealed that the simulated results of E were rated significantly less different from the measurement than all other software. Simulations of software A performed significantly worse than all others.

4.6 Discussion and Summary

In contrast to the previous investigations which only investigated room acoustic parameters, the analysis of a variety of simple and complex scenarios based on

(room) impulse responses presented in this chapter allows better insight into room simulations and their drawbacks. When discussing the results of this round robin investigation, it should be considered that participants were uninformed about the measurement results and they were – with the exception of the reflection on the diffuser of scene 1 – not allowed to adjust or modify the input data. However, this can not be regarded as a strict rule as the applied simulation tools use different interfaces for the input of the scene data. This potentially lead to participants eventually using slightly different data for the sound source directivity (either impulse responses or one-third octave energy spectra), the receiver directivity (e.g., different HRTF lengths), the scattering coefficients or the geometry. Such modifications (and simplifications) should therefore be regarded as part of the modeling process. Another general aspect, which should be considered during the evaluation, is the quantity of simulations each participant had to run. A simulation task additionally requires the import of the scene data, the configuration of the simulation and a review of the results. This explains why some of the submitted results might not have been properly reviewed before submission – and in some cases – also contain mistakes such as wrong source positions (scene 11), wrong source orientations (scene 3) or the simulations excluding diffraction in case of scene 5. Such issues, however, were also observed for less comprehensive scenarios of the previous round robin investigations [176, 29, 30, 31]. Bork reported that users failed to activate air absorption or applied false room volumes [31]. The sheer amount of data which had to generated in the current round robin also explains the lower number of participants compared to the previous investigations. The accuracy of the provided input data is another, potentially the most relevant, factor which should be considered when analyzing and interpreting the results of the participants. For the simple scenes the effect of uncertain input data is minimized considering the frequency range the data was analyzed for (cf. Fig. 3.32 in Section 3.8.2). Despite the measurement uncertainty of this data being low, the validity of the input data remains uncertain as even the application of normal- or random-incidence absorption coefficients instead of angle-dependent complex impedances can lead to relevant differences in a simulation result [5, 32].
For the complex room scenes, the provided absorption data is only indirectly based on measurements and although the geometries of the rooms were accurately determined, the level of detail was based on a decision made by the organization team. Thus, the definition of the input data involves subjective decisions, which ideally *should not be part of a prediction model* [46]. However, these subjective decisions were taken by an expert group and by instructing the participants, not to modify the common input parameters, the individual decision of each participant with respect to the input data modeling was reduced to a minimum. While one could expect, that a fitting of input data leads to less deviations to

the measurements and also among the participants' result, the application of a common dataset allows a comparison of different simulations and avoids that shortcomings of the simulation models are compensated by adjustments of the input data.

4.6.1 Reference scenes

Although scenes 1-5 are simple by definition, the focus on one acoustical effect requires a high modeling accuracy and corresponding input data. Simulating a matching temporal and spectral response even of a simple situation can be more challenging than predicting the energy decay rate of a complex room. More options to model the sound source directivity, the sound propagation and sound reflection are reasons why the results in the time and frequency domain differ even for simple cases such as a single reflection on a surface. Based on this analysis, shortcomings of simulations strictly based on GA can easily be demonstrated, for example by evaluating the spectral deviations for reflections on a reflector panel of finite size (scene 2) or on a reflector array (scene 4). For theses scenes, especially for scene 5, which concerns diffraction, wave-based simulations are more adequate tools than GA-based simulations.

Scene 1 (RS1): Single reflection (infinite plate)

The single reflection on a quasi infinite reflector showed that even very simple situations lead to varying simulation results. While all participants were able to model a direct sound and a reflection arriving at the correct time, only some participants were able to simulate a result, which corresponds to the measurement in the frequency domain, and achieved this only in case of the reflection on the rigid floor and the absorber. Remarkably, not one participant simulated results which exactly matches the result of another contribution. This allows the conclusion, that each simulation, especially if the user is also considered as part of the simulation, is different from another and leads to at least to minor deviations even for very simple situations as the specular reflection on a rigid floor. The lack of accuracy of the simulation results for the reflection on the diffuser is surprising, especially as diffuse reflections in room simulations are considered to be an essential element [50, 48]. Although the typical modeling of diffuse reflections by using scattering coefficients applied to a smooth surface represents a simplification of the scene, more accurate simulations results were expected, even if the simulations do not apply enhanced image source models for rough surfaces [26].

Scene 2 (RS2): Single reflection & diffraction (finite plate)

The evaluation of this scene revealed a drawback when applying fundamental GA models such as the image source model, which assumes a reflection on an infinite surface [2], without accounting for the finite size of a reflector. Despite this effect is well-researched and can easily be modeled [138, 137], only one participant submitted results which considered the effects of the finite reflector panels. The result analysis, however, also showed that in situations including a finite reflector for a receiver position without a possible specular reflection, no simulation software was able to model a corresponding correct reflection. While reflecting panels are typically only applied in concert halls, similar effects apply for objects in rooms such as cupboards or tables.

Scene 3 (RS3): Multiple reflection (parallel finite plates)

In case of scene 3, the participants were able to accurately simulate the fluttering reflections between the two parallel walls. To fully model the situation correctly, it is also required to account for the limited size of the panels and to accurately consider the directivity of the sound source. This was achieved at least to some extent by two participants in this scene.

Scene 4 (RS4): Single reflection (reflector array)

For the reflection on the array it was demonstrated that a simple model for reflectors of infinite size is insufficient in more complex situations, and could even lead to more erroneous results for a certain frequency range. Such a model, potentially only based on one limiting frequency (cf. Section 3.3), was applied by participant *C*. More extensive models for modeling such reflections exists [160], but might be challenging to integrate in an efficient holistic room simulation software. In this scene, interestingly, participant A, who did not submit results for scene 2 and whose simulated data showed systematic errors throughout all other scenes, delivered results which did not accurately resemble the measured data, but provided the most robust results independent of the reflection points being on or between the elements of the array. The fact that only three of the six participants submitted data for scene 4 also indicates that the three participants did not find it worthwhile to submit results potentially knowing that their applied software has limited capabilities with respect to an array reflection.

Scene 5 (RS5): Diffraction (infinite wedge)

For similar reasons as for scene 4, only two submissions included reasonable results for scene 5. Most participants applied simulation models not accounting

for diffraction and thus decided not to participate in this scene. Both participants applied first-order diffraction in case of a blocked direct sound path and generated similar impulse responses. In the frequency domain, substantial deviations could be observed when compared to the measured reference which also included various diffracted sound paths. In literature various models have been presented [22, 171, 164], which most certainly would perform better than the two contributions in this round robin investigation. Yet again, these models are mostly not integrated in (available) room simulation software which could be used to generate results for a wide range of scenarios as in this investigation. Interestingly, a GA-based diffraction model [152] has also been implemented for the RAVEN simulation model [153], which was used by one of the participants. This experimental module, however, was not maintained during the ongoing development of the RAVEN software and eventually excluded from the simulation environment in the software version used in this round robin.
Despite the investigation of the simple scenes highlighted many flaws of GA-based simulations, such shortcomings might only have a minor effect when calculating impulse responses or room parameters of complex rooms. It can, however, be stated that if these errors occur in the early part, that most likely relevant coloration is introduced.

4.6.2 Complex scenes

In contrast to the accurate scene description for the first five scenes, the complex scenarios were defined using estimated acoustic surface properties with limited validity for the entire frequency range (cf. Fig. 3.32). Thus, differences between measurement and simulation may either stem from shortcomings of the simulation models or uncertainties and errors in the description of the boundary conditions. With respect to the discussion of the simulation results of the complex scenes, the reader should be reminded that this situation corresponds to the application of room simulation on typical consulting scenarios, where the acoustician has to plan the room acoustics of a future building.

Scene 8 (CR1): Coupled rooms

None of the simulation programs produced energy decay curves which resembled the ones of the measured coupled room situation. On the one hand, these deviations are influenced by erroneous boundary condition data, on the other hand, a deviating effect was increased due to inhomogeneous absorption distribution in the less reverberant room, which could not be correctly accounted for by the simulation programs. Additional reasons for the deviations include inadequate configuration of the simulation (e.g., number of ray tracing particles)

or insufficient simulation models excluding diffraction modeling, which is especially important with respect to the coupling from one to the other room. The effect of the boundary conditions could be quantified by comparing the measured reverberation time ratio with the reverberation times based on the Eyring equation and the provided input data. It can not be stated that the participating programs in general are not able to model double slopes in case of coupled room scenarios. In general, if sufficiently many rays are emitted, a ray tracing algorithm is able to model the energy decay of coupled rooms [100]. If, however, a small gap connects both rooms, as it is the case for the smaller door opening angle $\phi_1 = 4.1^\circ$, the applied models also have to consider effects such as diffraction and or sound transmission at these gaps [186]. Additional investigations are required to determine by much and using which method the input parameters have to be adjusted to generate simulation results resembling the measured double slope.

Scenes 9, 10 and 11 (CR2-CR4)

The arrival time analysis of all three scenes indicated, that, with a few exceptions, the participants correctly defined sound source and receiver positions. An analysis of the initial time delay gap of the simulated RIRs gave further insights if the geometry was correctly defined in the simulation environment of the round robin participants. A deeper analysis to check whether the sources and receivers were correctly oriented and the all material coefficients were correctly assigned was not conducted as the that kind of information is challenging or even impossible to retrieve from the simulated RIRs and BRIRs.

The evaluated room acoustic parameters show that the tested room simulation software lead to consistent results, especially if the software is based on the same fundamental concepts. A trend of decreasing variance among the algorithms with increasing frequency was observed for most parameters in all three rooms. When compared to the corresponding outcomes of the first and second round robin on room acoustical simulation [176, 29], the amount of result variation among the participants has clearly decreased. While this might have been expected considering that all algorithms are based on fundamental GA concepts, at least one algorithm produced outlying results. Interestingly, these outliers came from two different algorithms depending on the room and the acoustic parameter. In general, for an estimation of energy parameters without knowledge of the measured results, the simulated parameters for the mid frequency range of almost all participating software are mostly in an acceptable range, and in some cases, the mean values are within the subjective limen of the evaluated parameters, which has to be considered as a critical perception threshold.

While for scene 9 it should be stated that the lowest evaluated octave band is below the Schroeder frequency of the room and thus the applied simulation models, based on GA, are not valid. For the medium and large room (scene 10 and scene 11), however, the room volume is sufficiently high for the application of geometrical models for the full evaluated frequency range. Especially consistent deviations of all simulated T20 values for frequencies below 1 kHz in case of scene 10 (cf. Fig. 4.15) have to be attributed to inaccurate boundary conditions descriptions, impacted by complex structures and various coupled volumes in the chamber music hall. This thesis is supported by the evaluation of the Eyring reverberation time, which shows also a trend towards considerably higher reverberation times for the lower frequency bands. The average deviation of all simulated T20 values to the Eyring reverberation time is smaller than the average deviation to the measured values. With a few exceptions, simulated T20 values are consistently higher than the Eyring values, which matches observations made for simulations of non-diffuse rooms reported in other studies [51, 9].

Independent of the Schroeder frequencies of the rooms, largest deviations and highest variances among the participants are observed for the 125 Hz frequency band. While erroneous boundary conditions are the obvious explanation for deviations from the measurements, the variance among the simulation results has to be attributed to potentially minor, but relevant differences of the simulation software and related signal processing steps. Interestingly, similar findings were already reported in the previous round robin investigations on room simulation [31].

The temporal RIR analysis showed that the most prominent reflections are simulated, but a substantial amount of diffracted and scattered energy is missed by most simulations. This is expected as a 3D model in general contains less surfaces than the real room. Additionally, the participating software did not account for diffracted sound, e.g., originated by the edge of the stage or by the seating. The lack of this energy might, in some cases, be compensated by more energy in the later part of the RIR. This, however, leads to potentially audible coloration, especially as this affects the early part of the RIR long before the predicted perceptual mixing time [95] of the rooms.

The outcomes of the temporal and parameter analysis suggest that differences between simulation programs and measurements are audible. However, it should be kept in mind that the perceptual study is based on the same rooms, but on different source/receiver positions and characteristics, requiring more signal processing tasks than the RIR and parameter calculation. This difference could lead

to more deviations from the measurements, but could also mask shortcomings of the simulations. The spectral analysis of the BRIRs (cf. Fig. 4.14), however, also revealed substantial spectral deviations well above theoretically defined difference limen. These differences were detected in the *authenticity* experiment, where all four softwares could be clearly distinguished from the measured BRIRs. For the less strict *plausibility* criterion, the auralizations were rated as being plausible in case of three of the four participating softwares. In the more detailed experiment on the perceptual qualities, the strong systematic deviations in the spectrum of participant A were also detected as perceptual differences. Most simulations by B, C and E were rated as brighter than the measurement, but remarkably low-frequency overestimation and high-frequency underestimation of reverberation times by the simulation algorithms could not be confirmed by the perceptual investigation, even if deviations were substantially higher than the JNDs .

This round robin investigation demonstrated various shortcomings of GA-based room simulation. In general, several advanced models have been proposed in the scientific community, e.g., phased GA models [32] or hybrid simulation models [4, 65], which are able to improve the result of the prediction. It is however, challenging to combine and integrate such models in a software, to validate and evaluate these models for (most) potential application cases and to implement all required interfaces for importing the input data.

5

Informed simulation of reference scenes

The round robin investigation presented in Chapter 4 has demonstrated that acoustical simulation results are often not in agreement with the corresponding reference measurement. On the one hand deviations can be attributed to inadequate simulation models, on the other hand the deviations are attributed to inaccurate input data. The discussion of the input data in Section 3.8.2 showed that the most severe limitations are observed for the data of the surface materials. For rooms which are not yet built, the simulation process is similar to the procedure followed in the previously presented round robin. Thus simulated results of planned buildings have to be regarded with a high level of uncertainty. For rooms, however, which already exist and can be described by acoustical measurements, simulations can be repeated until they match the measured results. This is often done by manipulation of the absorption of the 3D model, as this is typically the most uncertain input data. Such process is part of the task of an acoustic consultant in case of a planned renovation or modification of an existing hall.
This chapter presents and discusses results generated using the RAVEN simulation software. In these simulations either different datasets of absorption coefficients are applied or different ways to model and describe the sound source or the scene geometry are investigated. Additionally, binaural simulations of some scenes are created and discussed with respect to audible differences of corresponding auralizations. As the measured results of the corresponding situations are known beforehand, such simulations are either called *a posteriori* or *informed* simulations. Examples of informed simulations can be found in various studies [133, 170], also for scenes of the BRAS database [161, 166, 65]

The relevant data for the informed simulation of the scenes discussed in this chapter, including the simulation project files, the room model data, directivity files and the boundary condition data is published as a research data publication [10].

5.1 Selection of input data for informed simulations

Mostly due to a lack of suitable input data describing the physical properties of the boundary conditions, the definition of the room simulation's input data involves subjective design decisions [46] and is far from following a standardized procedure. The definition of the room geometry, scattering and absorption coefficients creates a multidimensional space with all of these datasets having an impact on simulated room acoustical parameters.

When preparing a simulation of a known space, the room model has to be be defined and initial coefficients for the surface absorption and scattering have to be determined or selected from databases [128] or literature [46, 177]. Following this procedure, the user is for instance able to model the acoustical effect of room elements either by detailed geometry and lower scattering coefficients or by simplified geometry with higher scattering coefficients [180, 31]. Another simplification is the definition of manageable number of wall materials which are assigned to the polygons of the 3D model. Due to a lack of explicit rules for this modeling process, the user can be sure that the simulation does not lead to correct results, often exceeding the JNDs of the simulated parameters [178, 14]. This is also a subjective process, and it can be expected that results would differ from user to user, especially if the users have a varying expertise and experience using room simulation models [178].

As high levels of uncertainties have been shown for standardized measurement techniques of the absorption coefficients [84, 63, 173], operators of room acoustic simulations typically choose to adjust the defined absorption coefficients so they would better correspond to the measured values in case of existing spaces.

5.2 Calibration of the boundary conditions

The *calibration* procedure corresponds to the process of adjusting input data of a room simulation until it matches or resembles one or more result parameter of a corresponding room acoustical measurement. As the calibration process is based on what is typically the output of a simulation, this process is often also described as an *inverse* approach [109]. In literature, various approaches to systematically adjust input data for room simulations have been suggested. Benedetto and Spagnolo [21] proposed the application of a ray tracing model to improve the determination of absorption coefficients for reverberation chamber measurements in case of non ideally diffuse sound field conditions. Pelzer et al. [125] presented an optimization model to match simulated and measured energy histograms for

the RAVEN simulation software. In the ODEON software, a calibration tool was integrated which applies a genetic algorithm to find optimized absorption coefficients of the room's materials [45]. Pilch [130] also uses a genetic algorithm, more specifically a differential evolution algorithm to find an optimized set of absorption coefficients. Here, the open-source software i-Simpa [129] was extended and as optimization parameters, multiple ISO 3382 parameters can be selected. Results, however, showed that it was challenging to achieve deviations of less than one JND if multiple room parameters were optimized simultaneously. A systematical calibration of the input data of the CATT acoustics software with respect to auralizations was presented and evaluated in listening experiments by Postma and Katz [132].

5.2.1 Target parameter: Reverberation time

With the reverberation time typically being the most important parameter when discussing the acoustic properties of a room [88], this parameter is the obvious choice when a simulation is matched to measured values. If the energy in a room decays exponentially at the same rate throughout the decay process, a correct simulated reverberation time would also lead to other energy parameters such as the early decay time, clarity or definition being similar to the corresponding measured values. As real rooms, however, do not establish ideal diffuse fields [88] and often show non-exponential energy decays, e.g., caused by inhomogeneously distributed absorbing material [91] or by coupled volumes [47], the derivation of parameters based on the simplifying concept of the reverberation time should be carefully considered.

For the RAVEN simulation environment, a possibility to adjust the absorption coefficients in order to match the simulated results to a target reverberation time is included in the MATLAB interface, which is part of the ITA-Toolbox [25]. The user is able to call the method *adjustAbsorptionToMatchReverbTime* of the class *itaRavenProject* repeatedly for a defined number of iterations. The method adjusts the frequency dependent absorption coefficients $\alpha_{i+1}(f)$ by multiplication of the current absorption coefficients $\alpha_i(f)$ with the factor $k(f)$:

$$\alpha_{i+1}(f) = k(f) \cdot \alpha_i(f) \tag{5.1}$$

The factor $k(f)$ is calculated according to Eq. (5.2).

$$k(f) = \frac{1 - (1 - \bar{\alpha}_i(f))^{(T_i(f)/T_G(f))}}{\bar{\alpha}_i(f)} \tag{5.2}$$

with $T_i(f)$ being the current simulated frequency dependent reverberation time (T30), evaluated and averaged for all sound source and receiver positions. $T_G(f)$

is the target reverberation time the user can define, either for 31 one-third octave or for ten octave frequency bands values (from 20 Hz to 20 kHz). The average absorption coefficient of the current iteration $\bar{\alpha}_i(f)$ is obtained by evaluating the equivalent absorption area according to the Eyring equation:

$$\bar{\alpha}_i(f) = \frac{1}{S}\left(-S * log\left(1 - \frac{1}{S}\sum_{n}^{N} S_n \alpha_n(f)\right) + 4\, m(f)\, V\right) \tag{5.3}$$

with S being the total surface area and V the volume of the room, N the number of materials and $\alpha_n(f)$ and S_n the absorption coefficients and the surface area covered by each material, respectively. The air attenuation coefficient $m(f)$ is calculated according to Bass, et al [19].
For most cases, after 5-10 iterations, the simulated reverberation times only show minimal deviations from the target values. It is beneficial if the initial absorption coefficients lead to a simulated reverberation time which does not substantially differ from the target reverberation time. With respect to the distribution of differently absorbing materials in the simulated room, it is also reasonable to start with values taken from the literature or reference measurements instead of completely arbitrary absorption values. Random starting values would also converge towards the target value for an adequate number of iterations, but would not maintain the relative differences of absorption in the room.

When discussing the quality of the calibrated data, it should be noted that this approach obtains absorption coefficients, for which only the simulated T30 values, averaged for all sound source and receiver positions, match the target reverberation time. This means that T30 values for individual positions not necessarily match the measurements of the identical position. Another issue is the filter synthesis of simulated room impulse responses. The described calibration method evaluates T30 values ($T_i(f)$) based on simulated energy histograms and not on room impulse responses. Signal processing and stochastic operations of the applied filter synthesis step can thus introduce additional deviations from the target values. The mid range and higher frequencies are only slightly effected by this, while the lower frequencies are more prone to relevant deviations. The difference of evaluated reverberation times based on energy histograms and RIRs is also discussed in subsection 5.3.1. The results of the reverberation time matching process for scene 9 and scene 11 are presented in the corresponding subsection of Section 5.3

Impact of the scattering coefficient

Most calibration models focus on the adjustment of absorption coefficients of one or all materials of the room. Accurate diffusion or scattering coefficients, however, are also challenging to obtain and despite having a lower impact on the simulated energy decay process, can also be considered for the calibration. Already in 1996, Lam [90] discussed different scattering implementations and the role of the input data. Later, Embrechts [51] extensively investigated the link between the energy decay and the scattering coefficient. An analytical model was developed which describes how and when the reverberation time calculated by simulation models is affected by the scattering coefficient. A similar investigation was conducted for the RAVEN software, which lead to an empirical model for the calculation of the reverberation time in dependency of the scattering coefficient (Eq. (5.4)) [9].

$$T(\alpha, s) = T_{s=1} \cdot (1 + k \cdot e^{\frac{-2s}{\alpha}}), \tag{5.4}$$

With α being the absorption coefficient, s the scattering coefficient of the room and $T_{s=1}$ the reverberation time simulated for fully diffusing surfaces ($s = 1$), corresponding to the maximum possible isotropy of reflections and, in case of uniform absorption distribution in the room, matching the Eyring reverberation time equation. The room factor k depends on the volume, the surface area, the shape and the structural complexity of the room. The formulation of the exact relationship for this is subject of ongoing research.
Simulated reverberation times[1] compared to the model predictions for different rooms, investigated for scattering coefficients ranging from 0 to 0.4 are presented in Fig. 5.1. To all three room models, only one surface material with $\alpha = 0.05$ was assigned, leading to a homogeneous absorption distribution in the rooms. It is observed that lower scattering coefficients lead to higher reverberation times in case of all three rooms, and that the prediction model resembles the simulated values. The room factor k, however, was chosen manually. When comparing the rooms, it is observed that the factor k decreases with decreasing complexity of the room geometry. Also, for these examples, the difference ΔT between reverberation times for $s = 0$ and $s = 0.4$ also decreases (0.93 s for the shoebox room, 0.57 s for scene 9 and 0.49 s for scene 11). This can be explained by introducing more geometrical scattering when a room is modeled by more details, less parallel surfaces and therefore also less regular reflection patterns for ray tracing particles.

[1] The simulations are conducted with the RAVEN software for a maximum image source order of 2 and for 50,000 ray tracing particles per octave band, diffuse rain technique was enabled.

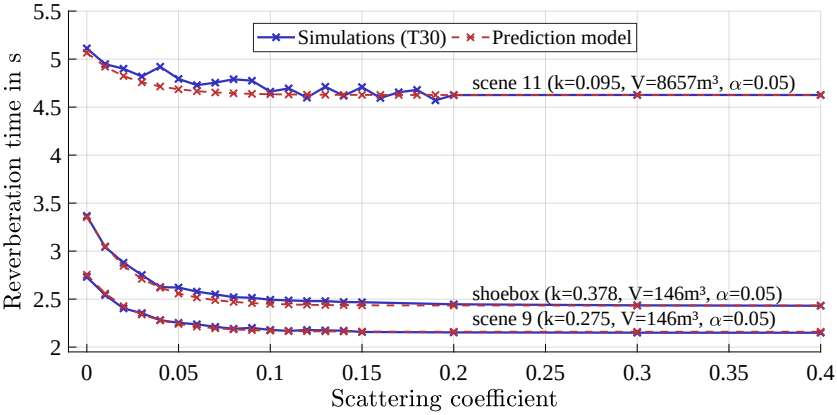


Figure 5.1: Simulated and predicted reverberation times for three rooms

5.2.2 Target parameter: Energy histograms

For many rooms, non-exponential decays, or if evaluated in the logarithmic domain, non-linear energy decays can be observed, e.g., for non-uniform absorption distribution or for coupled room situations. In this case, the calibration of input data using the reverberation time as a target parameter can lead to substantial errors when other room parameters are considered. To overcome this issue, Pelzer et al. [125] proposed a concept to match simulated and measured energy histograms. This was implemented in a MATLAB tool using the RAVEN simulation environment in the default hybrid simulation configuration combining image sources and energy histograms obtained by a ray tracing algorithm.
The numerical inversion of the simulation process is achieved by an iterative least-mean squares optimization. The cost function $f_c(\alpha)$ is described in Eq. (5.5) [125].

$$f_c(\alpha) = E_{RT}(t,\alpha) + E_{IS}(t,\alpha) - E_{meas}(t) \tag{5.5}$$

With α being a vector containing the absorption coefficients of all room materials, E_{RT} the simulated energy histogram based on the ray tracing results, E_{IS} the simulated energy histogram based on the image source model and E_{meas} the measured energy histogram. The minimization of the cost function is achieved using non-linear least-mean squares optimization. This is implemented very efficiently by calculation of the energy histograms using matrix multiplications. For this, the simulation results of RAVEN are exported as plane wave lists (image sources) and as wall hit logs (ray tracing), independent of the absorption coefficient of the room surfaces. This approach is suitable to match the measured and

simulated energy histograms, but faces the same issues as the reverberation time matching approach when individual positions and impulse responses instead of energy histograms are considered [14].

With the perspective of using this approach for the determination of boundary conditions of real rooms, evaluations have shown that it is possible to retrieve absorption coefficients of multiple materials in irregular distributions in case of hidden reference simulations, but only if the simulated scenario exactly corresponds to the reference with respect to geometry, scattering and source and receiver positions [6]. If the situations do not match, substantial deviations for the resulting set of absorption coefficients are observed. This corresponds to the situation if a reference measurement is applied instead of a reference situation, as room geometry, scattering and source and receiver positions cannot be determined with a very low level of uncertainty. Thus this approach is not suitable for the determination of boundary conditions of complex real rooms, it can, however, be applied for the determination of the absorption of materials in controlled laboratory measurements, potentially improving ISO 354 measurements, as suggested in the early publication of Benedetto and Spagnolo [21] and implemented by Probst [134].

5.3 Informed simulations of selected scenes

This section presents the informed simulation of four BRAS database scenes. All simulations are conducted using the RAVEN simulation environment (Version 2019.v3) for different sets of input parameters. Table 5.1 lists the four scenes and the variation of the input data. The focus of the selected scenes for the informed simulation is put on the complex room scenes and the corresponding boundary conditions because the discussion of the BRAS database in Chapter 3 showed that these scenes and the related input data have the highest uncertainties. This is supported by the analysis of the round robin investigation presented in Chapter 4, which revealed that for the room scenes, a substantial amount of the deviations are attributed to the boundary conditions of the complex scenes.

To investigate the effect of the (in case of higher frequencies) directional dodecahedron based RIR measurements, results of scene 9 are also simulated for varying sound source configurations. Scene 1 is investigated for the reflection on the diffuser as for this scene configuration many contributed results of the round robin failed to model a correct response in the frequency domain. Therefore, two simulations of the diffuse reflection using different approaches to model the diffuser are compared to each other and to the measurement.

Table 5.1: Overview of simulated scenes and input data variations

Scene	**Input data variation**	**Analyzed output**
Scene 1 (RS1)	Scene geometry and boundary conditions	RIRs for diffuser
Scene 8 (CR1)	Boundary conditions: Three datasets	RIRs
Scene 9 (CR2)	Boundary conditions: Four datasets Sound source directivities	RIRs & BRIRs
Scene 11 (CR4)	Boundary conditions: Four datasets	RIRs & BRIRs

5.3.1 Simulation model and configuration

The RAVEN environment allows the user control various configurations of the simulation. In the default setting, RAVEN applies a hybrid simulation model combining an image source model for the direct sound and the early reflections with a ray tracing algorithm which contributes the remaining part of the energy response. The transition between two models is defined by the maximum order of image sources, typically defined as 2. Thus, the image source model provides specular reflections up to second order, while the ray tracing model provides energy histograms which include early diffuse reflections (up to order 2) and all reflections of order 3 and higher. While the image source model outputs a list with audible image sources, the ray tracing model generates energy histogram for different frequency bands. Based on this, a filter synthesis process generates RIRs [151, 7, 16].
The most important parameters, separately presented for image sources, ray tracing and filter synthesis, are listed below.

- Image sources (IS)
  - Spectral resolution: 1/3 octave bands (31 frequency bands)
  - Maximum reflection order N_{IS} (default value: 2)
- Ray tracing (RT)
  - Spectral res.: 1/1 or 1/3 octave bands (10 or 31 frequency bands)
  - Number of particles per frequency band N
  - Maximum energy loss per particle ΔL_{max}
  - Radius of the detection sphere r_d

  - Length of a timeslot Δt
- Filter (room impulse responses) synthesis
  - Spectral resolution: Inherited from ray tracing configuration
  - Maximum reflection density Nd_{max} (Default value: 20,000 1/s)

Sound source directivities are applied as one-third octave magnitude spectra, in post processing the on-axis free field impulse response of the sound source is applied to the simulated RIR by convolution. The latency introduced by the on-axis impulse response is corrected by means of temporal shifting. In case of the BRIR-simulation, HRTFs of a length of 192 samples are applied for the direct sound, for early specular reflections and also for the reverberation tail. The directional resolution of the detection sphere (cf. [151] and [7] for details) in case of the ray tracing is set to 5° for both, azimuth and elevation.
If not stated otherwise, the presented configurations, and the various other parameters which can be set in the project file, are set to their default values.

Stochastic uncertainty

The simulation involves many deterministic but also some stochastic processes. The calculation of the image sources and the corresponding filter synthesis of the room impulse response containing the direct sound and early reflections is purely deterministic, while in the ray tracing model, the decision if a particle is being scattered or reflected specularly is based on a random number process [177], depending of the scattering coefficient of the corresponding wall material. Another variation can be introduced if the number of ray tracing particles is varied as this changes the launching angles of the particles.
The standard deviation σ_{EH} of the energy histogram obtained with the ray tracing model is theoretically described by Eq. (5.6) [175].

$$\sigma_{\mathrm{EH}} = 4.34\sqrt{\frac{V}{N\,\pi\,r_d^2\,c\,\Delta t}} \tag{5.6}$$

with V being the volume and c the speed of sound. The other variables correspond to the ray tracing configuration described in the list above. This equation describes the standard deviation of an energy multiplication-based ray tracing and assumes that the number of particles and the number of detected hits is constant throughout the entire simulation process. This also leads to the standard deviation of the energy histogram being independent of the time.
In case of results simulated with RAVEN various factors cause to a different stochastic behavior:

- The energy histogram also includes deterministic image sources
- The effect of the (deterministic) diffuse rain technique
- Maximum energy loss of particles (time variant particle count)

These factors introduce variations of the standard deviation over time. In Fig. 5.2, the standard deviation (SD) of 100 simulations for scene 9 (LS1–MP1) is presented. Due to the frequency dependent absorption and scattering coefficients, the SD also varies substantially for the six one-third octave bands. The chosen simulation configuration (N=10^5, r_d=0.8 m, ΔL_{max}=75 dB, Δt=2 ms) results in a calculated theoretical standard deviation of σ_{EH} = 0.14 dB, using Eq. (5.6). For the early part of the energy histogram (< 250 ms), this corresponds to the SD of the simulated results in all six depicted frequency bands as visualized in Fig. C.3 in Appendix C.
For the SD in the later part of the energy histogram, it is observed that for the 125 Hz and 200 Hz frequency band, values exceed 2 dB around the T30 value. This

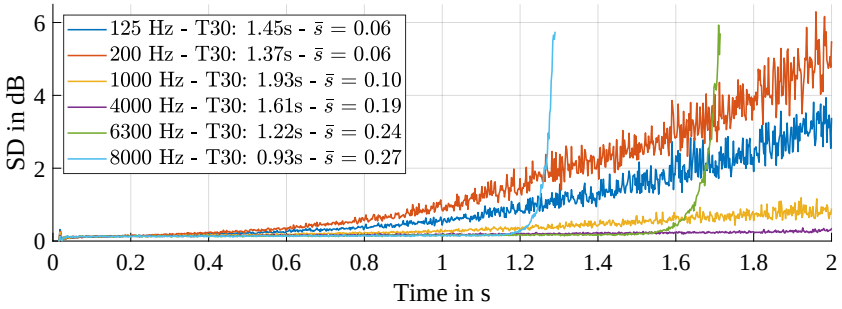


Figure 5.2: Standard deviation (SD) of 100 simulated energy histograms for scene 9. Six one-third octave frequency bands are shown. Simulation configuration: N=10^5, r_d=0.8 m, ΔL_{max}=75 dB, Δt=2 ms (σ_{EH} = 0.14 dB). $\bar{s}$ is the average scattering coefficient of the room.

is not observed for the higher frequency bands, despite the average absorption being higher (and reverberation times being lower for 6.3 kHz and 8 kHz). In general, higher reverberation times and higher average room scattering coefficients $\bar{s}$ lead to lower SD values. In case of the 6.3 kHz and 8 kHz frequency bands, the SD value rapidly increases at 1.2 s and 1.6 s, respectively. At these points in time, occurring shortly after the corresponding reverberation times, the ray tracing particles vanish as they reach the maximum energy loss level of 75 dB. These

points are shifted to later points (and generally all SD values for the later part are reduced) if ΔL_{max} is increased to higher values (cf. Fig. C.4 in Appendix C for ΔL_{max}=90 dB).

A second random process is included in the filter synthesis based on the energy histograms. As the reflection density in a room increases over time with t^2 [47], for each sample of the RIR it is decided if a reflection is set or not, based on a random process which depends on the current calculated reflection density. This ensures that the RIR does not contain any temporally regular reflections and therefore avoids the introduction of coloration during the filter synthesis. Fig. 5.3 presents the effect of the stochastic process in the ray tracing and in the filter synthesis on T30 values. While the reverberation time based on the energy histograms has consistently low SD values throughout all frequency bands (maximum SD: 0.009 s for 400 Hz), the SD and deviations from the T30 values based on the energy histograms are increased for frequencies below 300 Hz with a maximum SD of 0.24 s at 20 Hz. The mean of RIR-based T30 values exceed the upper JND threshold of the energy histogram T30 values only for some frequency bands below 300 Hz. The subjective difference limen for these low frequencies, should, however, only be considered as a reference value as no experimental evidence is available, that the JNDs are valid for the entire audible frequency range.
These uncertainties caused by stochastic processes should be considered in general when the results are evaluated. RAVEN also provides the option to generate deterministic data using the flags *fixReflectionPattern* and *setFixPoissonSequence*, fixing the random number sequences for the scattering and the filter synthesis calculation. This would remove statistical variation of multiple simulations, but for each investigated simulation it would be required to check the deviations of the fixed random number sequence.

5.3.2 Scene 1 (RS1): Single reflection (infinite plate)

This section presents the informed simulation results of scene 1 for the reflection on the diffuser. The simulations are conducted for two different scenarios. In the first scene setup, the diffuser is modeled by its exact geometry, consisting of rectangular profiles with alternating heights of 12 cm and 24 cm. The absorption and scattering coefficients of the material *mat_MDF12mm_plane_00deg*, taken from the BRAS database, is assigned to the surface of the diffuser. In the second scene setup, the diffuser is modeled by a flat surface with a height of 24 cm. To this surface, the absorption coefficients of the same material are assigned, but scattering coefficients are calculated according to Eq. (3.10) (see Section 3.7.2) for a characteristic depth d_{char}=0.12 m. The side view of both scenarios including

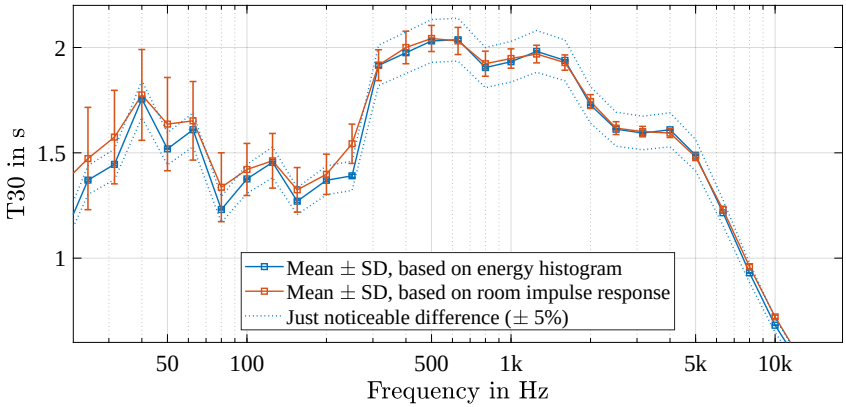


Figure 5.3: Evaluated T30 values (scene 9), averaged for 100 simulations, based on energy histograms and room impulse responses. Simulation configuration: N=10^5, r_d=0.8 m, ΔL_{max}=75 dB, Δt=2 ms. Ray tracing/Filter resolution: 1/3 octave bands. Whiskers show the standard deviation (mean value ± SD).

all sound source and receiver positions is presented in Fig. 5.4.

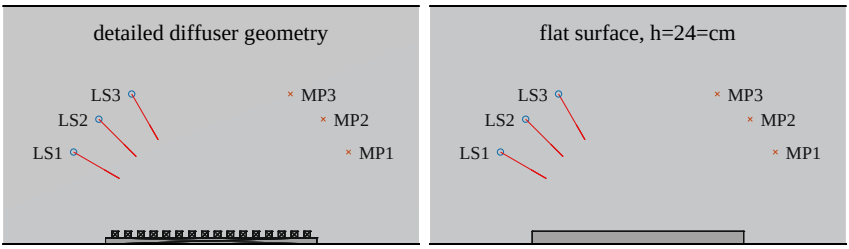


Figure 5.4: Two different input datasets for the simulation of the reflection on the diffuser in scene 1. On the left, the diffuser is modeled by its detailed geometry, on the right, a flat surface with higher assigned scattering coefficients is applied.

Because RAVEN simulations require a closed room geometry (if the ray tracing model is applied), the hemi anechoic chamber of the reference measurement is also modeled. The absorbing surface is defined by an absorption coefficient of 0.60

for 20 Hz, which increases linearly until it reaches 0.995 for 315 Hz. This room is indicated by the outer grey box in Fig. 3.10. For the flat surface, the question arises which height the reflector should have. Here, the reasonable choices lie within 12 to 24 cm. As the results showed best temporal alignment with the measurement for a height of 24 cm, most results are analyzed for this value.

Simulation configuration

RAVEN is primarily intended for closed room simulations, thus, for the simulation of a single reflection in a hemi anechoic environment, the configuration of the simulation has to be adjusted compared to a typical room situation. The maximum image source order was set to N_{IS} =1 as neither the plane surface nor the detailed diffuser structure would create any relevant audible image sources of higher order. To achieve a sufficient temporal and spatial resolution for the ray tracing, the time slot length was set to Δt =3 ms and the radius of the detection sphere was set to r_d =0.2 m. The number of particles was set to N =200,000 to ensure that every profile of the diffuser contributes diffuse rain to the energy histogram. The filter synthesis based on the ray tracing's energy histogram is done for full octave frequency bands. During this filter synthesis, an invalid assumption of the simulation becomes apparent. To determine the number of reflections in the synthesized impulse response, the reflection density N_r is calculated using Eq. (5.7).

$$N_r(t) = 4\pi \frac{c^3 t^2}{V} \qquad [s^{-1}] \tag{5.7}$$

This equation, proposed by [47], would ideally require a diffuse field and at least a closed room, which is not fulfilled for the anechoic situation of this scene[2]. Because the filter synthesis of the ray tracing requires a value for the room volume, the room was also modeled with almost totally absorbing surfaces of the anechoic chamber (see above), leading to a room volume of V =295.5 m^3. For the arrival time of the first specular reflection at 18 ms (in case of LS2–MP2), the evaluation of Eq. (5.7) results in $N_r(t) = 556\,\mathrm{s}^{-1}$, which corresponds to 0.013 reflections per sample and, on average, to 1.7 reflections per timeslot (Δt=3 ms). Thus, it is likely that at least one diffuse reflection is added to the impulse response during this time, which is, despite not being based on the valid physical model, a reasonable value.

[2] A figure showing the calculated reflection density for scene 9, 10 and 11 is shown in Fig. C.2 in the Appendix

Simulation results

Results are presented in the time and frequency domain for two RIRs, LS2–MP2 and LS1–MP3. Figure 5.5 compares the spectrum of the measurement, the simulation of the detailed diffuser model and of the flat surface/box with a height of 24 cm. For LS2–MP2 both simulated curves do not resemble the measured frequency response, but show an irregular comb filter structure. The simulated results also substantially deviate from each other, which is not the case for LS1–MP3. Here the simulated results are almost identical, the resemblance with the measured curve is also greater. The analysis of the results in the time domain,

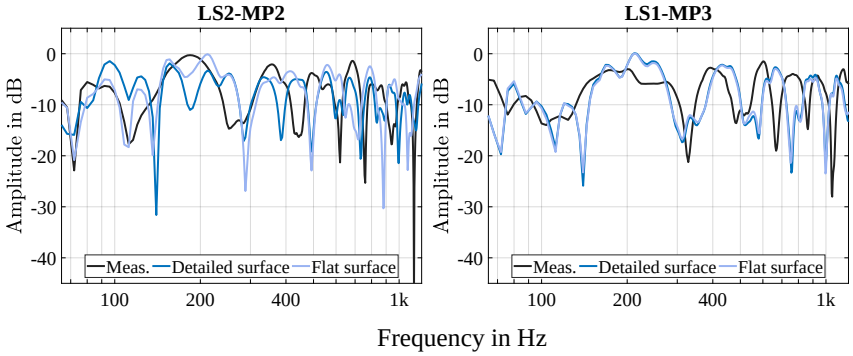


Figure 5.5: Frequency responses of measured and simulated results for scene 1, source/receiver combination LS2–MP2 and LS1–MP3

which are presented in Fig. 5.6, gives more insight in why the frequency responses of the two simulations differed for LS2–MP2. In the symmetrical situation of LS2–MP2, no first order image source was found if the detailed diffuser model was used. Interestingly, the remaining scattered energy is almost identical and thus, for these cases, independent of the geometry of the diffuser. The time domain results in general, also for LS1–MP3 (see Appendix, Fig. C.1), show a good visual resemblance of measurement and simulation, especially if the diffuser was modeled by a flat surface. In all cases, the simulated direct sound, which also shows the highest amplitudes, matches the measured results.
Independent of the diffuser modeling and the configuration of the simulation, deviations of the applied simulation model are always to be expected, introduced by the simplifying approach of the GA model. Due to the limited complexity of this scene, a BEM simulation can easily be applied up to the mid-frequency range and also compared to the GA simulation and the reference measurement.

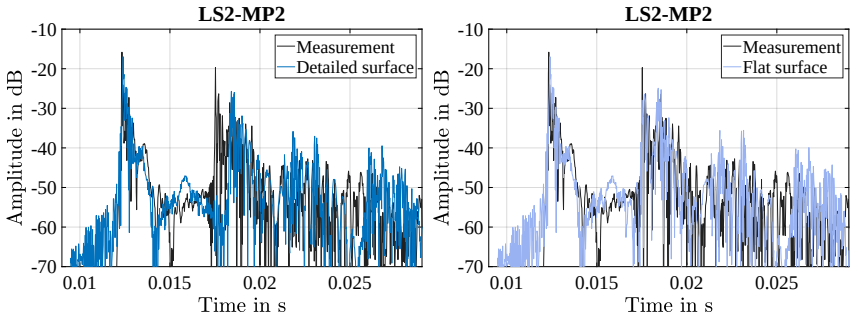


Figure 5.6: Impulse responses for scene 1, measured and simulated (detailed and flat surface) for LS2–MP2

For LS2–MP2, the results in the frequency domain up to 1 kHz are presented in Fig. 5.7. In this example, the GA simulation is based on a diffuser modeled as a box with a height of 12 cm, because for this setup the resemblance in the frequency domain was higher than for a height of 24 cm. As expected, the result based on the BEM model, which was simulated with the software COMSOL, is very similar to the measured frequency response, but also does not align perfectly with the measured result.

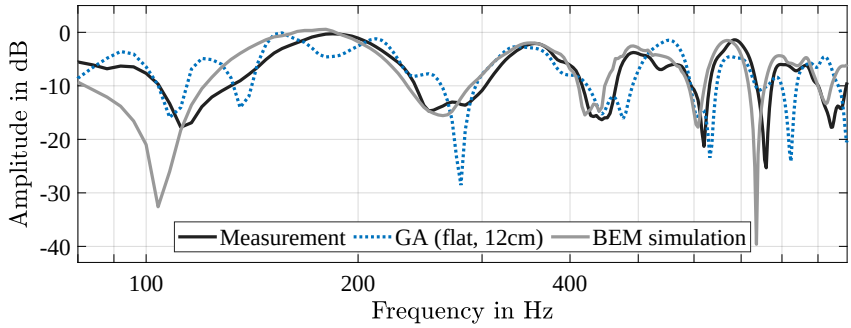


Figure 5.7: Frequency responses of scene 1 for LS2–MP2. Measurement, simulated with RAVEN for the flat reflector surface and simulated using the boundary element method (as included in the COMSOL software).

5.3.3 Scene 8 (CR1): Coupled rooms

Parts of this section have been published in: Aspöck, L., and Vorländer, M. (2019). Simulation of a coupled room scenario based on geometrical acoustics simulation models. In Proceedings of Meetings on Acoustics (Vol. 36, No. 1, p. 015002). Acoustical Society of America [17].

The results of the round robin investigation led to the question in how far GA-based simulation software are able to accurately predict a coupled room situation. In general, it has been shown in different studies, that GA-based simulation software can accurately simulate similar scenarios [100, 186]. Nevertheless, no contribution to the round robin delivered decent results. In fairness to the participants and in comparison to the other studies, it has to be mentioned that the round robin participants were not informed about the measurement results and had to simulate several other scenes, which potentially lead to a reduced effort with respect to the configuration of the simulation and the review of the results. To get more insight in the reasons for the poor performance of the simulation programs, more simulations using the RAVEN software were calculated for three input datasets including varying boundary conditions:

I. Initial coefficients of BRAS database
(correspond to the round robin input data for scene 8)

II. Fitted coefficients of BRAS database

III. Adjusted coefficients: Absorption and scattering coefficients are iteratively adjusted to match the decay rates of the measured coupled room scenario.

The absorption and scattering coefficients of all datasets are listed in the Appendix, in Table C.1. The different datasets can be characterized by the ratios of the calculated reverberation time using the Eyring equation of the individual rooms (r_{Ey}), which are listed in the upper part of Table 5.2.

Simulation configuration

All simulation results presented in this section were generated with the RAVEN software, applying a hybrid model using second order image sources and a ray tracing simulation including the diffuse rain technique for $N = 10^6$ particles. The maximum allowed energy loss for each particle was set to ΔL_{max} =75 dB.

Simulation results

Before the coupled room situation was simulated, both rooms were simulated separately (door angle $\phi = 0°$) for all three input datasets. The resulting T20

reverberation times and the corresponding ratio (r_{GA}) of these T20 values are listed in the lower part of Table 5.2. While the results of reverberation chamber (R_2) are similar to the reverberation times based on the Eyring equation, the reverberation times calculated by RAVEN are considerably higher for the laboratory room (R_1). This deviation, caused by inhomogeneous absorption distribution in R_1, also leads to consistently smaller ratios for the RAVEN simulations in case of all three datasets.

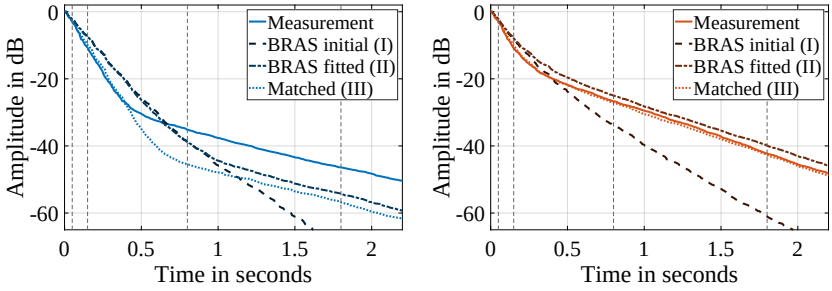
Table 5.2: Reverberation times calculated for uncoupled rooms R_1 and R_2 based on the three input datasets. r_{Ey} and r_{GA} describe the ratio of the reverberation times of both rooms based on the Eyring equation and the GA simulation (RAVEN), respectively.

	initial (I)	**fitted (II)**	**matched (III)**
$T_{\mathrm{Ey,\,R2}}$ [s]	2.49	5.84	5.42
$T_{\mathrm{Ey,\,R1}}$ [s]	0.79	0.79	0.63
r_{Ey}	3.15	7.39	8.60
$T20_{\mathrm{GA,\,R2}}$ [s]	2.50	5.76	5.41
$T20_{\mathrm{GA,\,R1}}$ [s]	1.30	1.25	0.83
r_{GA}	1.92	4.61	6.52

For the coupled room situation, the EDCs for LS2–MP3 were calculated using the ITA-Toolbox method *ita_roomacoustics_edc* after applying a 1 kHz bandpass. The EDCs for both door angles are shown in Fig. 5.8, in comparison to the measured EDCs. The simulation output using *initial* coefficients (dataset I) reveals significant deviations in both, the early and the late section of the EDC, and resembles most results of the round robin (cf. Fig. 4.10 in Section 4.4), showing only a weak double slope. For dataset II the double slope becomes more pronounced as the ratio of the reverberation times is more than two times higher than for dataset I. The early decay rate of the simulations using dataset II, however, remains lower than the measured curve despite the calibration of absorption coefficients according to the Eyring equation. This deviation can be explained by evaluating the uncoupled room simulations of R_1 using absorption coefficients of dataset II. For this scenario, the Eyring reverberation time matches the measured reverberation time, but the RAVEN simulations using dataset I or dataset II lead to substantially higher reverberation times (cf. Table 5.2). In the coupled situation, this results in a high value of DT_{early} for both door angles using dataset II. The processed decay times[3] of DT_{early} and DT_{late} for all scenarios are

[3] cf. Eq. (3.6) in Section 3.3

listed in Table 5.3. A considerably better resemblance with the measured EDC can be observed for the late decay of the dataset II simulations, especially for ϕ_B. This was expected as the uncoupled reverberation time $T20_{GA,\,R2}$ is close to the measured reverberation time of R_2.



(a) $\phi_A = 4.1°$ (b) $\phi_B = 30.4°$

Figure 5.8: Measured and simulated energy decay curves using RAVEN, evaluated for the 1 kHz octave band for LS2–MP3. Simulations were conducted for all three input datasets I-III. Dashed vertical lines indicate the early and the late time frame for the decay time evaluation.

Table 5.3: Calculated decay times DT_{early} and DT_{late} based on the energy decay curves obtained by measurement and by informed simulations using the RAVEN software

		Meas.	**Initial (I)**	**Fitted (II)**	**Matched (III)**
$\phi_A = 4.1°$	DT_{early} [s]	0.73	1.10	1.09	0.83
	DT_{late} [s]	5.29	1.96	4.21	5.45
$\phi_B = 30.4°$	DT_{early} [s]	0.74	1.01	1.07	0.85
	DT_{late} [s]	3.85	2.21	4.08	3.91

The definition of the input parameters of dataset III are based on an iterative manual adjustment of the absorption coefficients until the simulated EDC resembled the measured EDC, for the opening angle ϕ_B. This adjustment led to a high reverberation time ratio $r_{III} = 8.60$, but to a lower ratio $r_{GA,\,III} = 6.52$ (cf. Table 5.2) when considering the uncoupled room simulations using RAVEN. Interestingly, the uncoupled reverberation times of dataset III are very close to the actual measured reverberation times in the uncoupled situation.

Although Fig. 5.8b suggests a very low deviation for this scenario, the relative deviation of $\mathrm{DT}_{\mathrm{early}}$ is 13% while $\mathrm{DT}_{\mathrm{late}}$ is very similar with a relative deviation of 2%. The application of this adjusted dataset for the other opening angle ϕ_A indicates a problem of this approach. Here, the decay rates are also similar to the measured curve, but the turning point is shifted by about 200 ms and 15 dB.

5.3.4 Scene 9 (CR2): Small room (seminar room)

Despite being assigned to the category of *complex* rooms, this scene's room has a fairly simple geometry and also a rather homogeneous distribution of absorption in the room. This makes the room less challenging to simulate, which was confirmed by the round robin results of this scene. Especially for the mid frequency range (500 Hz and 1 kHz), the results of the participants were fairly consistent and mostly showed very low deviations from the measured values. However, as substantial deviations were observed for the lowest frequency band (125 Hz) and the evaluation of the BRIRs also showed relevant spectral deviations, it is worthwhile to conduct and analyze more simulations of this scene.
The first part of this section considers different data for the boundary conditions. Simulations are presented for the round robin input data, the two input datasets provided in the BRAS database and for a fourth input dataset, which was created by iteratively adjusting absorption coefficients until the simulated T30 value matches the measured T30 values[4], using the method described in Section 5.2.1. Results of the iterative matching process are depicted in Appendix C, in Fig. C.5.
The second part of this section deals with the modeling of the *omnidirectional* sound source. Here simulations are conducted either without or with source directivity data for the dodecahedron speaker. For both parts, room acoustic parameters are evaluated. In case of the boundary condition investigations, the spectra of BRIRs are additionally analyzed, while for the directivity investigation, the spectra of simulated RIRs are compared.

Simulation configuration

The RAVEN simulation environment was applied as described in Section 5.3.1 in the configuration generating up to second order image sources and energy histograms obtained by the ray tracing algorithm for 31 one-third octave frequency bands. Specifically, the following configurations were used:

- RIR-Simulations: $N = 5 \cdot 10^5$, r_d=0.8 m, ΔL_{max}=75 dB, Δt=2 ms
- BRIR-Simulations: $N = 2 \cdot 10^5$, r_d=0.8 m, ΔL_{max}=80 dB, Δt=2 ms

[4] Here, the averaged T30 values of ten RIRs, both for simulated and measured data, are considered.

For both cases, the parameters correspond to the configuration chosen for the stochastic uncertainty investigation of Section 5.3.1 (which lead to a standard deviation $\sigma_{\mathrm{EH}} = 0.14\,\mathrm{dB}$), with the exception of the particle count per frequency band being higher for the simulations presented here. Thus, lower stochastic uncertainties can be considered.

Simulation results for different boundary conditions

Simulations for the four different sets of boundary conditions were conducted and the corresponding room parameters of the ten RIRs were processed in the same way as in the round robin evaluation. The averaged T20 values are presented in Fig. 5.9. The results of the round robin input data and the *initial* coefficients do not show substantial deviations from each other. This was expected as these input datasets hardly differ, the differences involve minor geometry changes and different scattering data. The higher scattering coefficients of the initial dataset lead to lower reverberation times for 500 Hz and 1 kHz. The simulations using the fitted coefficients of the BRAS database (based on Eyring equation matching, cf. Section 3.7.1) lead to T20 values which match the measured values for frequencies above 500 Hz and slightly exceed the values of around one {JND for lower frequencies. Thus, the fitted coefficients based on the Eyring equation already provided an adequate set of absorption coefficients for the RAVEN simulation. As expected, the coefficients obtained by the T30-matching hardly differ from the fitted coefficients for most frequency bands and lead to simulated T20 values which all lie within the perceptual difference limen of the measured values. The analysis of each individual position with respect to the evaluated T20 values reveals that highest positional variations of the measured RIRs occur for the 250 Hz band (min: 1.43 s, max: 1.97 s) followed by the 125 Hz frequency band. The simulations, independent of the used input data set, show smaller positional variation. This is particularly striking for the 250 Hz octave band, where the T20 value for the matched input dataset is practically independent of the position (min: 1.67 s, max: 1.69 s). The corresponding figure including the minimum and maximum values is presented in the Appendix C, in Fig. C.6.
As described in Subsection 5.2.2, the calibration of one room acoustic parameter does not automatically lead to low deviations of other parameters. One example, the clarity value C80 for position LS1–MP5, is presented in Fig. 5.10. Here, it can be seen that the fitted and T30-matched dataset improve the situation for the higher frequency bands (4 kHz and 8 kHz) and lead to C80 values which correspond to the measured values. Nonetheless, for the lowest two frequency bands (125 Hz and 250 Hz), deviations of more than 3 dB are observed. For these frequency bands neither the measurement, nor the simulations show strictly linear

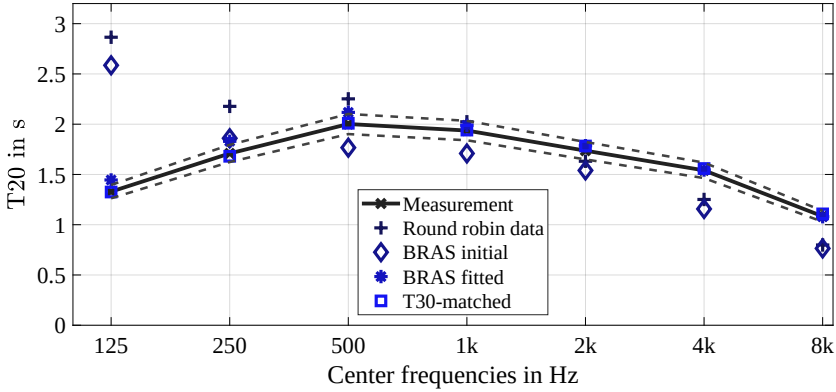


Figure 5.9: Evaluated T20 values (average of 10 positions) for simulated and measured room impulse responses of scene 9. Simulations were conducted with RAVEN for different input data sets.

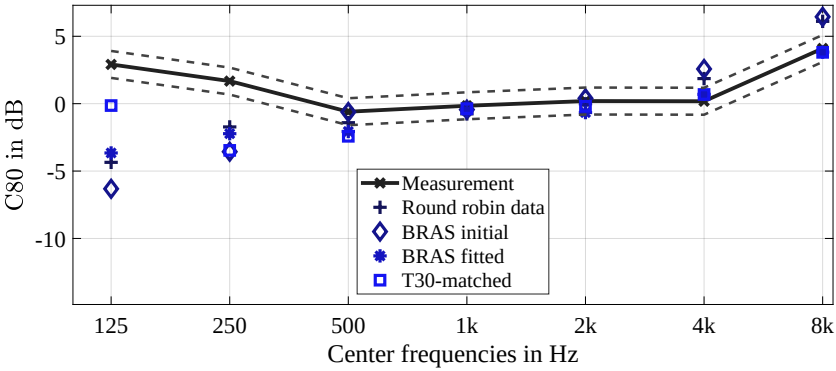


Figure 5.10: Evaluated C80 values for simulated and measured room impulse responses of scene 9, positions LS1–MP5. Simulations were conducted with RAVEN for different input data sets.

energy decays (cf. Fig. C.7 in the Appendix), which contributes to the different clarity values for these frequency bands. This highlights the importance of regular energy decays when the reverberation time is used as the matching parameter.

For the simulation of the BRIRs, two aspects should be taken into account when analyzing the results. First, the simulation process also includes a receiver directivity, which might introduce additional deviations when comparing measurement and simulation. Secondly, the fitted and the T30-matched coefficients are based on the averaged T30 values of the measured RIRs instead of the BRIRs. The energy decays at the position used for the BRIR simulation and measurements slightly differ from the energy decays of the RIRs. In how far these factors have a relevant impact on the simulation results can be analyzed when comparing the magnitude spectra of the simulations for different boundary conditions. The results in the frequency domain of one BRIR example, the left channel of LS7–MP6, is presented in Fig. 5.11. To all results in this figure, a one-third octave smoothing filter was applied.

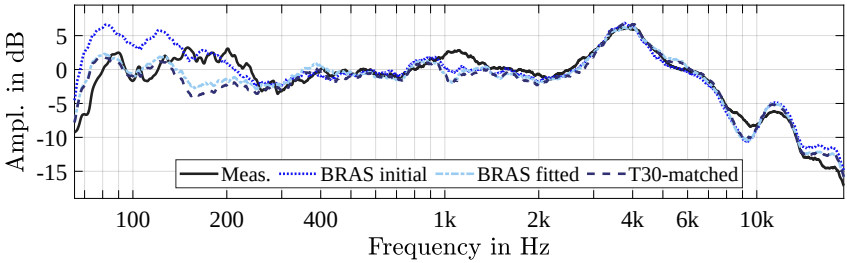


Figure 5.11: Magnitude spectrum of measured and simulated binaural room impulse responses, left channel, for position LS7–MP6

Above 300 Hz, the spectra of all three investigated simulations are similar to the measured spectrum. Differences of around 3 dB are visible at around 1 kHz and 9 kHz. For the lower frequencies, differences of up to 4-5 dB are observed for all three simulations. The increased energy around 100 Hz in case of the initial coefficients correspond to the overestimated reverberation time which was observed in the parameter analysis in the 125 Hz frequency band (cf. Fig. 5.9). However, there is less agreement for other BRIR positions. The spectral differences of two BRIRs are depicted in Fig. 5.12. To these results, also a one-third octave smoothing filter was applied before the difference was calculated. In general, the right channel showed higher deviations than the left channel. Especially for lower frequencies, deviations up to around 10 dB at 70 Hz occur for all three simulations in case of LS4–MP6. At 12 kHz all simulations show a 5 dB boost, which is also observed for the left channel, but is not visible in the right channel of the measured BRIRs.

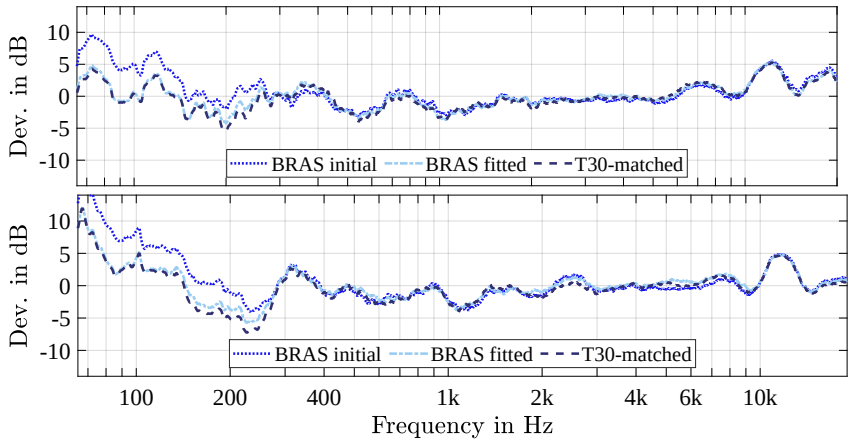


Figure 5.12: Deviation of measured and simulated spectrum of binaural room impulse responses, right channel, for position LS7–MP6 (top) and LS4–MP6 (bottom)

As BRIRs are primarily intended for the creation of auralizations, it is particularly interesting in how far the deviations are audible. First informal listening tests of the author have shown that it is worthwhile to assess the plausibility and the authenticity of the informed simulations. While for convolutions with pulsed noise bursts in critical listening conditions, differences between simulation and measurement could be detected, for speech or music samples, however, it was very challenging to distinguish the auralizations, especially for the center position. Example auralizations are included in the corresponding research data publication [10]. Using the results of the informed simulations, it would be of interest to repeat the round robin listening experiments with respect to authenticity and plausibility (cf. Section 4.5).

Simulation results for different sound source modeling

The dodecahedron measurement speaker which was used in the reference measurements of the single-channel RIRs introduces uncertainties as the speaker deviates from a perfectly omnidirectional radiation pattern. In this section, it is investigated in how far the consideration of directivity data and individual positions of the three-way loudspeaker (cf. Fig. 3.25 in Section 3.4) are relevant for the result analysis. Three different sound source configurations are considered:

I. Sound source modeled as a perfectly omnidirectional point source positioned at the center of the mid-range unit

II. One sound source at the reference position of the center of the mid-range unit using the measured combined directivity

III. Three sound sources at the corresponding reference positions of subwoofer, mid-range speaker and high-frequency unit, each assigned with the corresponding directivity data (cf. Section 3.4).

All simulations are conducted for the RIR setup (two source positions and five receiver positions, cf. 3.19) using the T30-matched absorption coefficients. For configuration III, the total RIR result is obtained by summation of the RIRs for the three positions.
The simulation configuration was similar to the other informed scene 9 simulations ($N = 2 \cdot 10^5$, r_d=0.8 m, ΔL_{max}=80 dB, Δt=2 ms).

Figure 5.13 shows the spectrum of the measured and simulated RIRs. As expected, the room modes below the *Schroeder* frequency $f_S = 235$ Hz, which are evident in the measured result, are not matched by the simulations. For the higher frequencies, the simulations including the directivity (and the on-axis frequency response) of the simulation resemble the measured spectrum, also showing a dip between 1 kHz and 2 kHz, which, however, is not as pronounced as for the measurement, especially for configuration II. In general, this example shows that

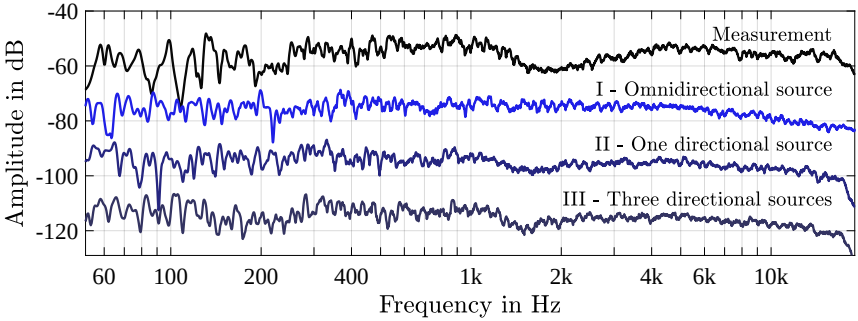


Figure 5.13: Magnitude spectrum of measured and three differently simulated room impulse responses for LS1–MP2. Simulated results are shifted by -20, -40 and -60 dB for better visibility.

if a spectral analysis is of interest, the directivity and the frequency response of the measurement speaker should be considered. The main intention of the

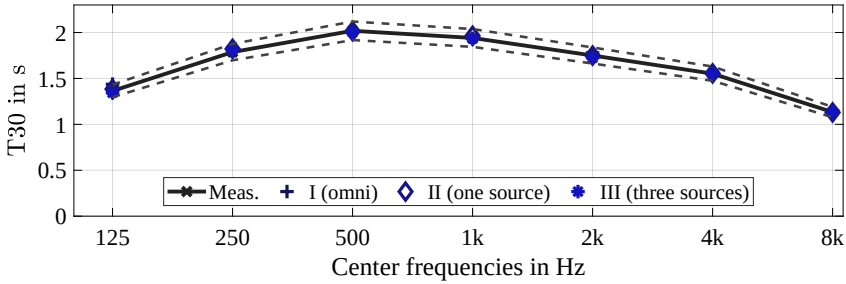


Figure 5.14: Evaluated T30 values (average of 10 positions) for simulated and measured room impulse responses of scene 9. The simulations are conducted for three sound sources configurations.

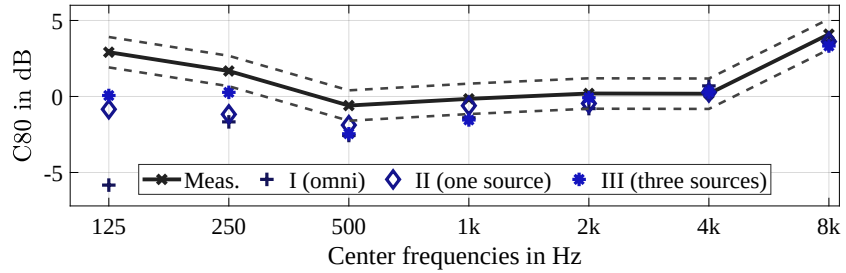


Figure 5.15: Evaluated C80 values for simulated and measured room impulse responses of scene 9, position LS1–MP5. The simulations are conducted for three sound sources configurations.

RIR measurements, however, is the evaluation of room acoustic parameters. The evaluated reverberation time T30 for the three sound source configurations and the measurement is shown in Fig. 5.14. Here it is observed that the evaluated T30 value is independent of the sound source modeling and matches the measured times for all three cases. For other parameters, such as the EDT, slight variations occur (see Fig. C.8 in Appendix C), but without showing a trend that the directional configuration leads to lower deviations. The C80 results of position LS1–MP5 depicted in Fig. 5.15 seem promising, but deviations in the 125 Hz frequency band are only slightly reduced on average across all positions can not barely be observed for the 250 Hz frequency band. In general this effect is rather unexpected as the applied directivity is practically omnidirectional for the 125 Hz

and 250 Hz frequency bands (cf. Fig. 3.27 in Section 3.4). More likely, this effect is attributed to the slightly different sound source position and the stochastic filter synthesis process (cf. Section 5.3.1), which leads to different results among configuration I, II and III. Larger deviations are also expected for parameters, which account for the total energy of the RIR instead of decay rates (T30, EDT) or energy ratios (C80). A deeper analysis of this, also including an evaluation of the strength parameter, is subject of ongoing research.

5.3.5 Scene 11 (CR4): Large room (auditorium)

Scene 11 is also investigated for the variation of the boundary condition datasets. In contrast to scene 9, the fan shaped room has a more complex geometry, a substantially greater volume and thus a lower *Schroeder* frequency $f_s = 31\,\text{Hz}$.

Simulation configuration

RAVEN simulations were conducted generating up to second order image sources and ray tracing based energy histograms for 31 one-third octave frequency bands. The following simulation configurations for RIR- and BRIR-simulation are applied:

- RIR-Simulations: $N = 2 \cdot 10^5$, r_d=0.8 m, ΔL_{max}=70 dB, Δt=2 ms
- BRIR-Simulations: $N = 2 \cdot 10^5$, r_d=0.8 m, ΔL_{max}=80 dB, Δt=2 ms

As the particle count N was not increased, the statistical error of the simulation is higher for scene 11 than for scene 9, due to its greater room volume. The theoretical standard deviation is $\sigma_{\text{EH}} = 0.77\,\text{dB}$, calculated using Eq. (5.6) for $V = 8657\,\text{m}^3$ and the simulation parameters listed above.

Simulation results for different boundary conditions

Simulations for four different boundary condition datasets were conducted. These datasets include the round robin data, the *initial* and the *fitted* coefficients of the BRAS database, and absorption coefficients obtained by the T30-matching process described in Section 5.2.1 for five iterations. The iteration process is documented in Fig. C.9, included in Appendix C. The scattering coefficients of the T30-matched condition are taken from the BRAS database, which are identical for the initial and fitted dataset.

The evaluated T20 values, averaged for ten RIRs, are shown in Fig. 5.16. Using the data of the round robin, substantial deviations from the measured value are observed for 125 Hz and 250 Hz, while the measured T20 values are slightly

underestimated for 2 kHz and above. The results of the initial coefficients are very similar to the ones of the round robin data, which was expected as the input data only slightly changed compared to the round robin data[5]. Simulations using the fitted coefficients based on the Eyring equation lead to lower deviations, but overestimate the measured values for 125 Hz, 250 Hz and for 1 kHz. The simulations for the T30-matched coefficients, as expected, produce T20 values which are very close to the measured data and only underestimate the measured reverberation time for 125 Hz, either caused by the fact that the matching was done for the T30 parameter or by the stochastic filter synthesis process.

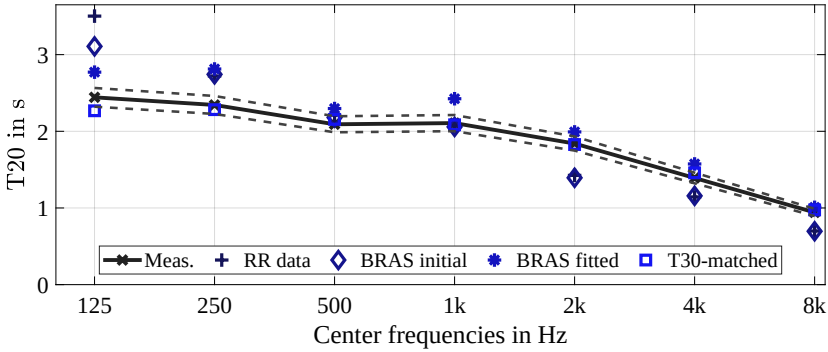


Figure 5.16: Evaluated T20 values (average of 10 positions) for simulated and measured room impulse responses of scene 11. Simulations were conducted with RAVEN for different input data sets.

While the evaluation of individual positions mostly shows good agreement between measurement and the simulated data, particularly using the T30-matched coefficients, for some position and parameters, remarkably high deviations are registered. One example is depicted in Fig. 5.17, showing clarity values for LS2–MP2. Here, the simulations consistently result in clarity values for frequencies below 2 kHz that are at least 1 dB, for 125 Hz and 250 Hz around 3-5 dB lower than the measured values.

To explain these deviations, the positions of the receiver and the sound source in the room have to be considered. MP2 is located close to the stage, in the front left part of the audience area, not very far away from the balcony and the wall on the

[5] In fact, only one material with less than 10 m^2 surface area was added to the model and all scattering coefficients were slightly modified

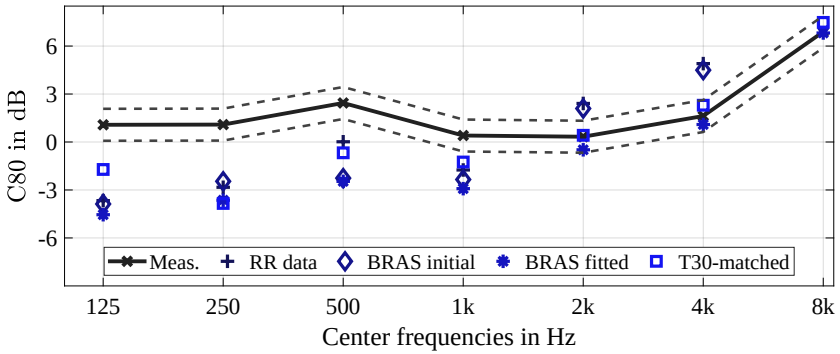


Figure 5.17: Evaluated C80 values for simulated and measured room impulse responses of scene 11, positions LS2–MP2. Simulations were conducted with RAVEN for different input data sets.

left side of the room (see Fig. A.5 in the Appendix). The distance to sound source LS2, which is located on the left part of the stage, is about seven meters. The vicinity to the side wall (5 m) and the balcony (3 m) leads to a substantial amount of reflected energy that is observed in the measured RIR, but not present in the simulated RIR, as visualized in Fig. 5.18[6]. While the direct sound, arriving shortly after 20 ms, coincides for measurement and simulation, the simulated RIR does not contain any reflections up to 40 ms. The simulated energy histogram based on the

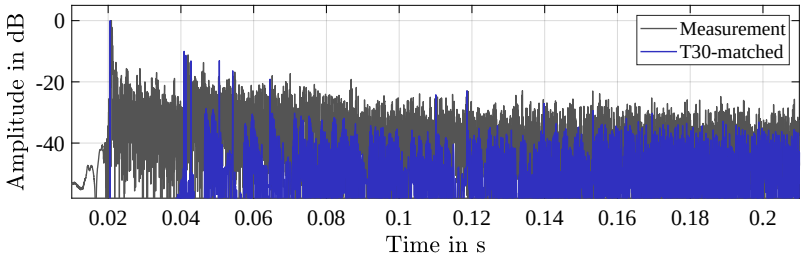


Figure 5.18: Early part of measured and simulated room impulse response for scene 11, LS2–MP2

[6] The modeling issue of the seating area, which is typically extremely simplified in simulation, was already noted in early validations of the ODEON software [110].

ray tracing contains energy during this time, but no reflections are inserted during the RIR-synthesis as the reflection density only increases very slowly due to the large room volume of the scene (cf. Fig. C.2). At 40 ms, the reflection density is lower than 100 reflections per second, the expected value of the corresponding random process accumulates to 1.25 reflections until this point in time. Thus, there is a considerable likelihood that no reflection is inserted. This represents a drawback of the simulation model, which can account for early scattered energy, but fails to accordingly synthesize the RIR. This can cause relevant deviations in situations with substantial amount of early, non-specular reflections in large rooms, as it is the case in the presented situation. The evaluated clarity values for the identical situation based on the energy histogram shows that the error is reduced by up to 2 dB (cf. Fig. C.10 in Appendix C). This demonstrates that, especially if the room simulation includes the synthesis of RIRs, achieving accurate results requires more than the adjustment of absorption coefficients. In an improved simulation model, it should be considered if the reflection density model should also be based on other parameters such as the ratio of specular and scattered energy in the early part of the energy response. Interestingly, this issue was already described in a publication by Lam, stating that models "*can only adequately approximate complex real-life effects when averaged over a sufficient number of reflections*" and thus "*are more likely to go wrong when predicting acoustic features that are highly dependent on the very early part of the impulse response where reflections are few and distinct*" [90].

The lack of this energy in the early part is also visible when comparing the simulated and measured BRIRs. The first 100 ms of the right channel BRIR of LS3–MP6 and the left channel BRIR of LS7–MP6 are presented in Fig. 5.19. For both BRIRs, a matching direct sound and floor reflection is observed and also the more prominent reflections coincide, but in general, the comparison reveals less energy in between the distinct reflections in case of the simulations. The effect however, is not as evident as for the RIR-simulation for LS2–MP2. The first reason for this is the geometrical surrounding of the binaural receiver position: MP6 is positioned in the center of the room, located on the aisle between the two main audience areas. The second reason is the distance to the sound source - with the direct sound arriving closely before 30 ms, the gap until the point in time when the ray tracing filter synthesis generates more reflections is smaller. Another difference, in comparison to the RIR results, is caused by the application of the HRIR filters (length: 4.4 ms, 192 samples @ 44.1 kHz), leading to more temporal smearing of the BRIR.

In the frequency domain, however, the results are dominated by the energy of the direct sound. One example, the left channel of the BRIR for LS7–MP6 smoothed with a one-third octave sliding window, is shown in Fig. 5.20. Due to the major

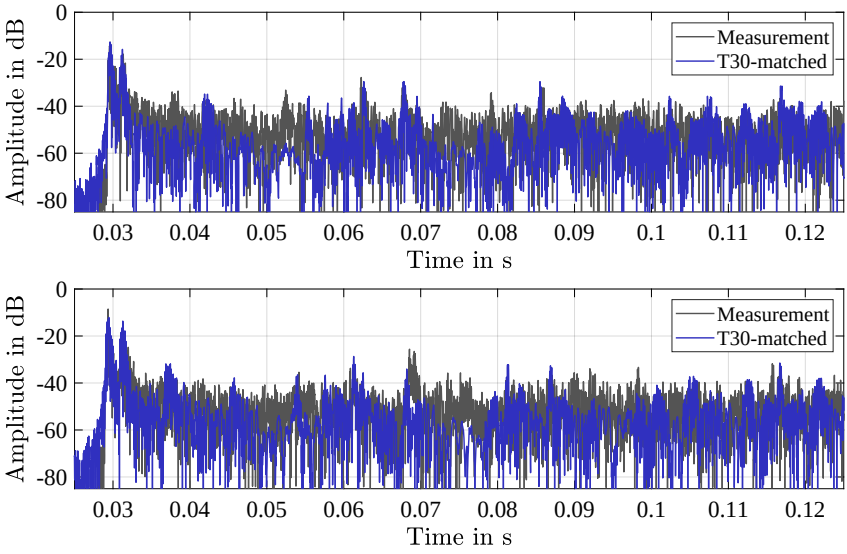


Figure 5.19: Simulated and measured binaural room impulse responses of scene 11. Right channel impulse response for LS3–MP6 (top) and left channel impulse response for LS7–MP6 (bottom).

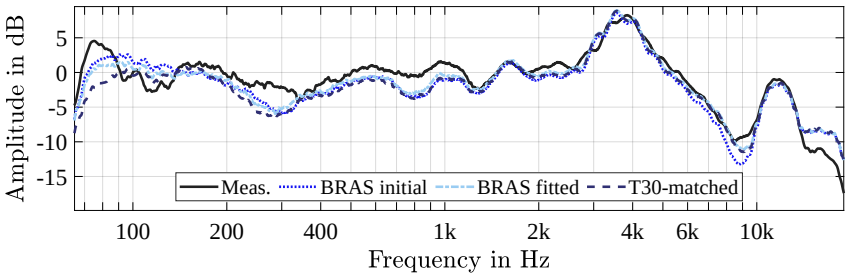


Figure 5.20: Magnitude spectrum of simulated and measured binaural room impulse responses of scene 11, LS7–MP6, left channel

impact of the direct sound, the results of the simulations using three different input datasets barely differ and mostly resemble the measured curve for frequencies between 1 kHz and 13 kHz. For lower frequencies, higher deviations are observed for all three simulations, with the highest differences being 5 dB and 4 dB for

75 Hz and 300 Hz, respectively. This suggests that audible differences could be detected when the simulation and measurements are compared in BRIR-based auralizations, in particular for critical listening conditions using pink noise bursts. A perceptual analysis of these informed simulations is subject of future research.

5.4 Discussion and Summary

For four scenes of the BRAS database the input data of the simulation was varied and informed simulations were compared to the corresponding measurements.

Results of the diffuser reflection (scene 1) have shown that in principle, different geometrical modeling approaches can be considered. It was demonstrated, however, that simple geometrical modeling in combination with the application of higher scattering coefficients lead to more reliable and consistent results with respect to the variation of sound source and receiver positions[7]. Both investigated options resulted in less deviations than the uninformed simulations presented in Chapter 4. Reasons for this are potentially a more scene-oriented configuration of the simulation and a more thorough review of the simulated results.

For scene 8, it was shown that GA-based tools are capable of simulating double sloped EDCs. Calibrated input data, however, does not overcome shortcomings of the simulation software which become evident, for example, in case of small aperture sizes. Calibrated simulations can produce correct results, but might show substantial deviations in other cases for parameters such as the turning point of the EDC. Nonetheless, it has to be kept in mind that the investigated scenario includes controlled extreme conditions with low practical relevance, especially concerning detailed parameters such as the turning point. Considering this, the performance of GA-based simulations for coupled room situations can be regarded as acceptable, if valid input data is available.

The simulations of scene 9 revealed that the fitted absorption data based on the Eyring equation lead to very low differences between simulated and measured reverberation time, also in the lowest evaluated frequency band. An additional calibration of the simulation according to the measured T30 values only had minimal effect on the applied absorption coefficients. A more detailed analysis of individual positions showed that, especially for more sensitive parameters such as clarity in the lower frequencies, substantial deviations can occur despite the input data calibration. These effects, although exceeding the perceptual limen

[7] Similar results were also found in the third round robin [31]

defined in the corresponding standards, might not be audible in auralizations. Preliminary informal listening experiments suggested that differences of auralizations based on the BRIR simulations using the T30-matched coefficients are challenging to identify. Thus, a perceptual investigation of these informed simulations is of interest and subject of future research.
The simulations using different data for the omnidirectional sound source showed that, with respect to the evaluated room acoustic parameters, the application of directivities for the sound source is not relevant for the simulated results. Only if a detailed spectral analysis is desired, the directivity and the frequency response of the loudspeaker should be considered.

In the analysis of scene 11, similar observations as for scene 9 were made. Here, the Eyring-fitted coefficients lead to some deviations for the simulated reverberation times, the T30-calibration lead to values within the perceptual difference limen of 5%. The repeated simulation using the round robin data resulted in values resembling the average of the results generated by the round robin participants. Analog to scene 9, some positions and parameters could be identified, which showed more discrepancies between simulation and measurement for the lower frequency bands. A deeper analysis revealed that these deviations can not be attributed to erroneous boundary conditions. Instead, they are caused by non-exponential energy decays, geometrical modeling and flaws of the simulation model, also with respect to the impulse response synthesis. When it comes to auralization and perception, these issues might be of less concern as the direct sound dominates the energy spectrum for the investigated positions.
In general, it is possible to conduct informed simulations which resemble measured data to a great extent. Still, for certain parameters, substantial differences can be registered, despite of the calibration of the simulation's input data. Some shortcomings of the applied T30-matching procedure can be overcome if an optimization of the total energy histogram is applied. Nevertheless, to achieve matching data not just for parameters, but also in time- and frequency-domain, the adjustment of boundary conditions is insufficient. Here, the applied simulation model has to be improved with respect to geometrical modeling (e.g., by application of level-of-detail concepts [33, 126]), reflection modeling and filter synthesis.

6

Conclusion and Outlook

After related work and the concepts of room acoustic measurements and simulations were presented, this thesis documented and discussed the BRAS database, a benchmark for room acoustic simulation that is intended for the validation of simulations. Based on this database, two studies involving different GA software were conducted and evaluated. In the first study, a round robin comparison was organized in which six participants submitted simulated data without prior knowledge of the scene's measured data. In the second study, room acoustic simulations of selected scenes were conducted by the author.

The BRAS database provides an extensive collection of room acoustical scenes ranging from a single reflection setup to a large lecture hall. The seven scenes of the database representing simple *reference scenes* are supplemented by four *complex rooms*. By including information about the scene geometry, boundary conditions, the source and receiver characteristics, and measured impulse responses for different source-to-receiver configurations, a user is able to check a simulation software for external validity. The provided measured data allows for a comparison and analysis of parameters, and in the time- and frequency domain. As seven scenes contain binaural receivers, simulations can be compared to the corresponding measurements in perceptual experiments.

Despite its restrictions with respect to the frequency range, which are mostly caused by limitations and uncertainties of the provided boundary condition data, the presented database is a versatile tool for the validation of acoustical simulations. For the reference scenes, geometry and boundary conditions are given with a high accuracy and therefore can be applied in validations at least for the frequency range from 100 Hz up to 4 kHz. In case of the complex rooms, the challenge to determine accurate absorption and scattering coefficients became evident. As the initially defined data, based on in-situ measurements and on tabulated values, was found to be inadequate for the input of GA simulations, a second set of absorption coefficients was added to the database using *fitted* values based on

the Eyring equation and the measured reverberation time values. These revised coefficients for the room scenes represent a more reasonable input data set for GA simulations, which was demonstrated for scene 9 and scene 11 in Chapter 5. Modeling the geometry of the complex rooms also introduces uncertainty to the simulation as it contains numerous modeling decisions, despite being partly based on accurate laser-aided measurements and the application of typical limits of 0.5 m for details. Thus, the creation of the 3D models is a subjective process leading to different 3D models if repeated by another person. Therefore validations of acoustical simulations using the complex scenes should always include a thorough analysis of the simulation results. Deviations of the simulation from the measurements can not necessarily be attributed to the simulation model, but help to identify shortcomings of the software (cf. simulations of scene 11 in Chapter 5).

The scope and the level of detail of the BRAS database is, to the author's knowledge, unique in the scientific field of acoustics. All related data is publicly available on a research data repository [13] and is open for potential future extensions, e.g., new scenes or more advanced measurement methods for acoustical boundary conditions.

Based on the developed database, a round robin comparison was conducted using five of the simple scenes and the four complex rooms as the test scenarios. Six participants using five different acoustic simulation algorithms submitted results which were then compared to the measured data. At this stage, the BRAS database did not include any measurements of the scenes yet. The participants thus conducted uninformed simulations, corresponding to the process of a simulation of a future building. As the participants were asked to submit impulse responses, more evaluation possibilities were possible in contrast to the first three round robins on room acoustic simulation. For three complex room scenes, also a round robin comparison using auralizations was conducted. In contrast to the previous round robins, the presented investigation does not only extensively document the deviations from measured results, but, with the help of temporal and spectral analysis, explains and highlights specific reasons for the deviations of the applied simulation software in more detail.
For future investigations, it would be interesting to analyze, in how far the results of radiosity models and simulations including higher order diffraction or wave-based approaches differ from the here applied GA models.

Two findings of the round robin investigation are particular interesting. First, the large deviations of some participants for the simple scenes. Admittedly, these scenes were specifically designed to evoke acoustical effects that require the consid-

eration of extended GA models in a simulation. Nevertheless, even the reflection on a one dimensional diffuser could not be accurately modeled by most participants. Secondly, the results of the complex rooms demonstrated yet again, that it is still challenging to determine adequate absorption and scattering coefficients for simulations if the user is uninformed about the real acoustics of the investigated room. The room parameter evaluation revealed, that simulated values only showed low deviations to the measurement in the mid-frequency range, especially in case of the reverberation time, but substantial deviations were found in the low and high frequencies, and for more sensitive parameters such as clarity. The analysis of these scenes in Chapter 5 showed that these deviations can be minimized to a great extent in the higher frequencies and to some extent in the low frequencies if the absorption data is adjusted.

If the user of a simulation knows acoustical details of a room or environment, e.g., based on acoustic measurements, this knowledge can be used to define the input data of the room simulation. This was discussed and investigated in Chapter 5 using different approaches, either based on the reverberation time or on the energy decay of the investigated room. Results have shown that for all four investigated scenarios, simulated data resembling the measured data could be generated. Excellent agreement was found in case of the reverberation time for scene 9 and scene 11. The simulated RIRs and BRIRs, however, still contain potentially relevant differences which became evident in both, the temporal and the spectral domain. Nevertheless, these deviations were challenging to detect in informal listening tests which encourages to conduct a perceptual study on the authenticity and plausibility of these simulations in future research.

Concerning the validity of a room simulation it should be noted that no GA-based simulation is a universally valid tool to describe the acoustic wave propagation in rooms as it is based on a simplifying approach replacing wave-based behavior by rays. Coming back to the statement "*absolute accuracy is at present some way off*" made by Mendel Kleiner in 1993 ([83], cf. Section 1), *absolute accuracy* with respect to reverberation times can be achieved[1], and potentially also with respect to the *authenticity* criterion [39] in perceptual evaluations of auralizations. Nevertheless, a simulated RIR based on GA models will never perfectly match a corresponding measurement or an analytical solution, in particular not for the full audible frequency range covering three orders of magnitude. Considering that in order to achieve these *accurate* results, the initially defined input data has to be adjusted, a general problem of room simulation validation for

[1] *Accuracy* shall be regarded as fulfilled if deviations are considerably smaller than the corresponding subjective difference limen.

complex rooms also becomes apparent, namely the lack of valid input data for the boundary conditions. Even if at some point in the future, an improved and standardized method for the acquisition of valid absorption coefficients or even complex impedances is available, the user of a GA-based simulation still has to define the adequate room geometry and/or scattering coefficients. For this part of input data, it is very unlikely that methods will emerge which will accurately determine this data in complex rooms. Therefore, the definition of input data for GA simulations will remain a subjective process [46] and relevant uncertainties must be accounted for in simulation results [178].

Nonetheless, with the intention of reducing these uncertainties of simulation results, validations of (future) simulation models and software still need to be conducted. With the BRAS database presented in this thesis, it is possible to validate new or existing simulations step by step for the different included acoustical phenomena. Here, it should be considered to add a new scene to the database, which bridges the gap between simple reference scenes and the complex rooms. Such a scene could include a simple, very controlled laboratory room including well-defined boundary conditions and (flexible) room elements to stepwise increase the complexity of the room and its acoustical effects. The design of such a standardized controllable room could be adopted from similar test rooms of previous studies [172, 122] and from existing rooms such as the "*Espace to projection*" at the IRCAM in France [113].

A

Additional BRAS data

A.1 Scene descriptions

Scene 4 (RS4): Single reflection (reflector array)

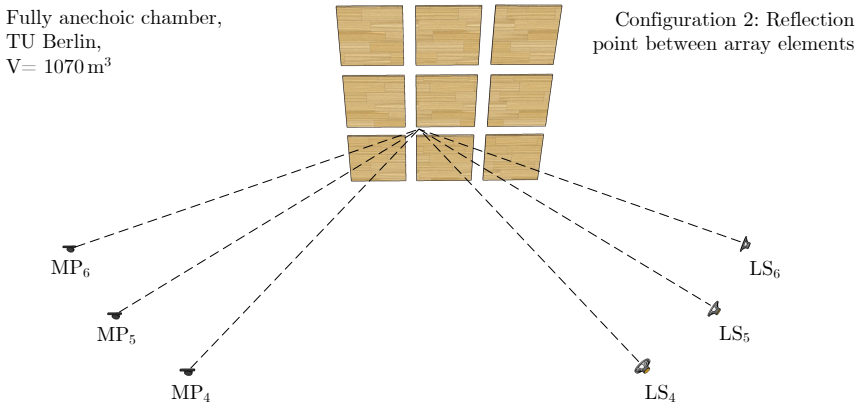


Figure A.1: Side view of scene 4 for configuration 2 including three sound source positions and three receiver positions

Scene 9 (CR2): Small room (seminar room)

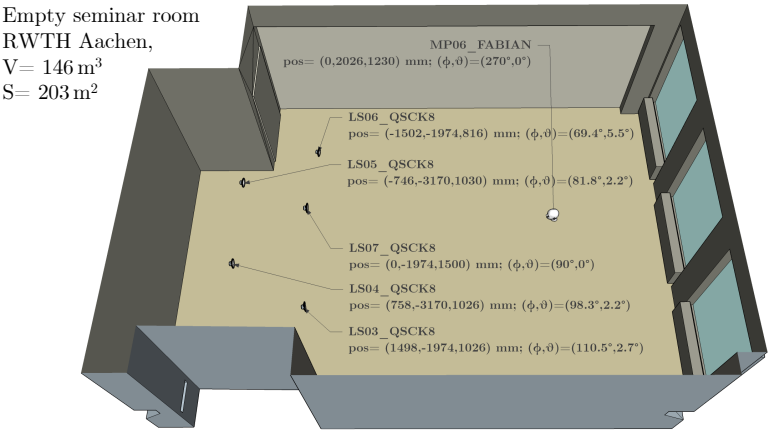


Figure A.2: Top view of scene 9, binaural setup for auralization including five sound sources and one receiver

Scene 10 (CR3): Medium room (chamber music hall)

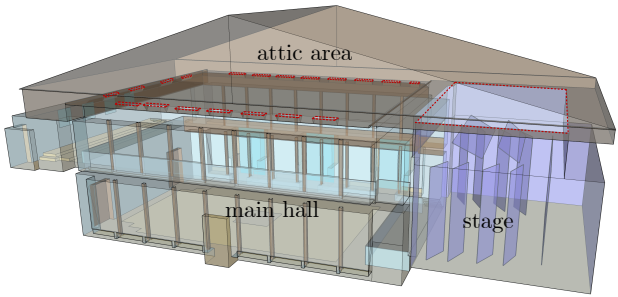


Figure A.3: Side view of scene 10 including the main hall, the stage and the attic. Red dashed lines indicate apertures connecting main hall and stage to the attic area.

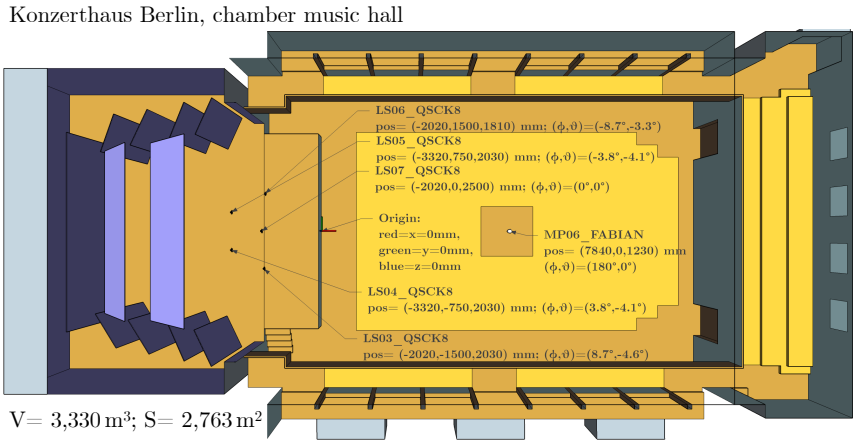


Figure A.4: Top view of scene 10, binaural setup for auralization including five sound sources and one receiver

Scene 11 (CR4): Large room (auditorium)

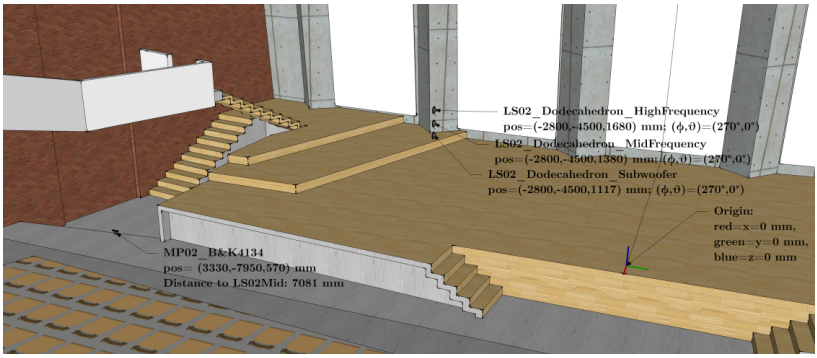


Figure A.5: Room view of scene 11 including sound source position LS2 and receiver position MP2

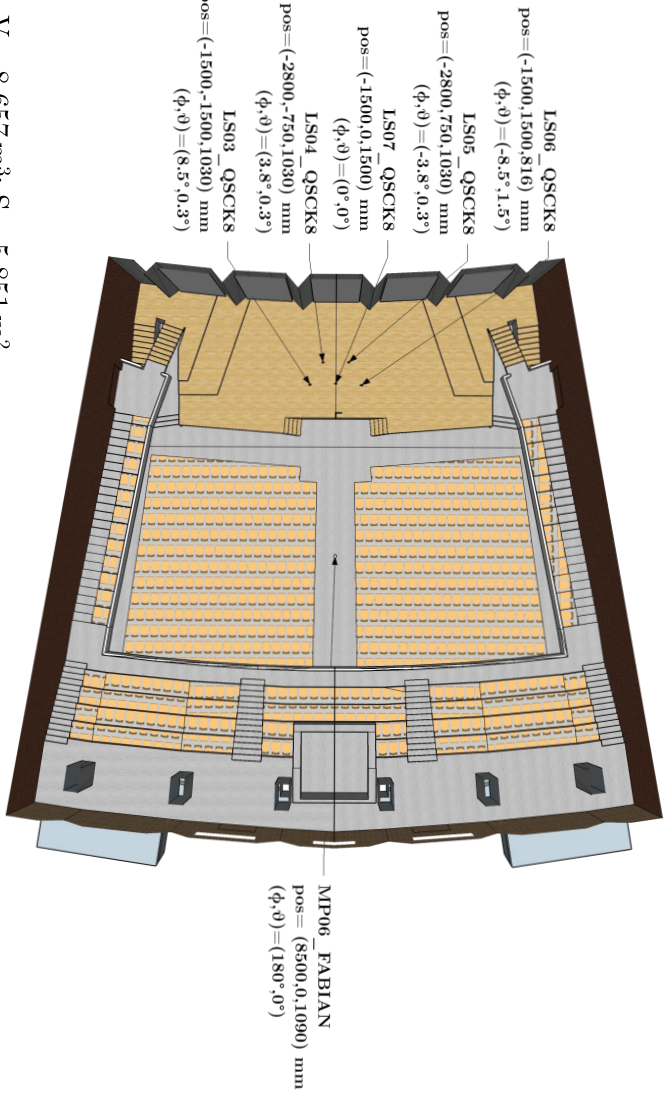


Figure A.6: Top view of scene 11, binaural setup for auralization including five sound sources and one receiver

A.2 Uncertainty and deviations in measured data

Table A.1: Estimated uncertainties and deviations for measured data of the BRAS database

	Uncertainty/ Deviation	**Details**
Combined general uncertainty	$< 0.53\,\mathrm{dB}$ ($f \leq 2\,\mathrm{kHz}$)	RIR measurement uncertainty determined according to Witew [188]
Scene 1 (RIRs)	$< 0.78\,\mathrm{dB}$ ($f \leq 16\,\mathrm{kHz}$)	Based on mean absolute deviations of corresponding time-windowed direct sound measurements for different scene configurations in frequency domain (1/3 octave bands, 18 RIRs).
Scene 9 (RIRs)	$< 0.75\,\mathrm{dB}$ ($f \leq 12.5\,\mathrm{kHz}$)	Based on mean absolute deviations of 10 repeated RIR measurements (three days later).
Sound source directivities	$< 0.4\,\mathrm{dB}$ ($f \leq 10\,\mathrm{kHz}$)	Spherical harmonics processing (interpolation) of source directivity.
Sound sources: Angular error	$< 0.56\,\mathrm{dB}$ ($f \leq 8\,\mathrm{kHz}$)	Mean absolute error in horizontal plane of Genelec 8020c caused by an assumed angular deviation of 3° (1/3 octave bands, 360 filters).
Sound sources: Angular error	$< 0.92\,\mathrm{dB}$ ($f \leq 8\,\mathrm{kHz}$)	Mean absolute error in horizontal plane of Genelec 8020c caused by an assumed angular deviation of 5° (1/3 octave bands, 360 filters).

B

Additional round robin results

B.1 Simple scenes

Scene 1 (RS1): Single reflection (infinite plate)

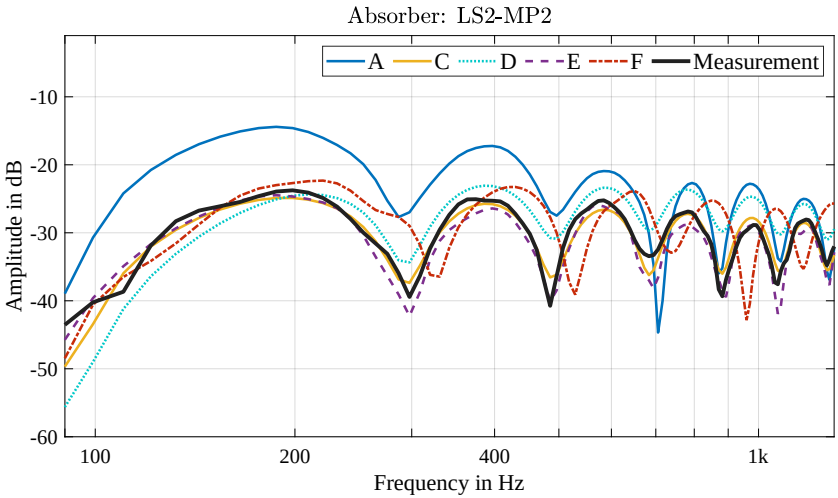


Figure B.1: Measured and simulated impulse responses of scene 1, reflection on absorber (4 m x 4.2 m) for LS2–MP2

Scene 2 (RS2): Single reflection & diffraction (finite plate)

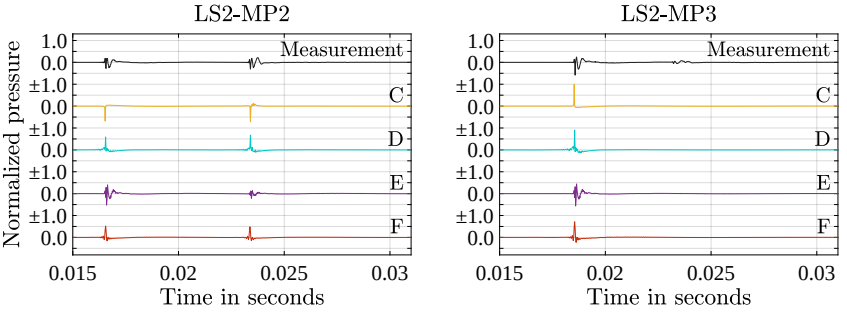


Figure B.2: Measured and simulated impulse responses of scene 2, reflection on rigid plate (1 m x 1 m) for two sound source – receiver combinations

Scene 3 (RS3): Multiple reflection (parallel finite plates)

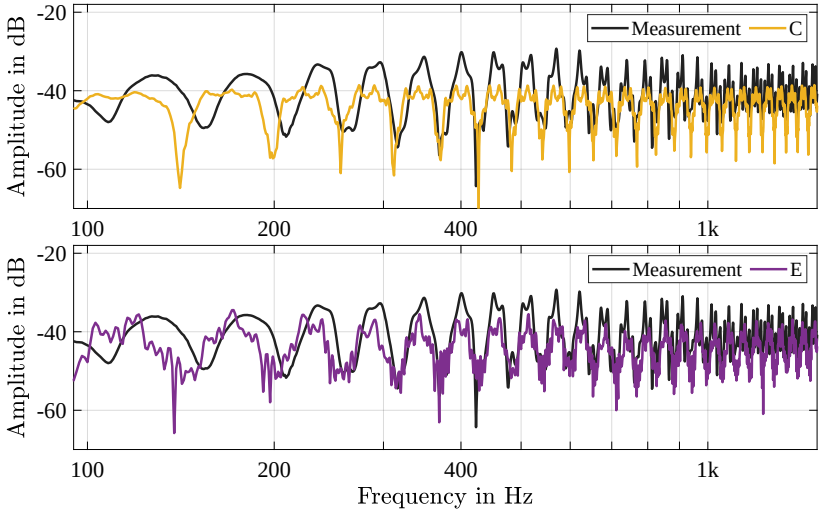


Figure B.3: Measured and simulated spectra of scene 3 for LS1–MP1. Only results of participants C and E are shown.

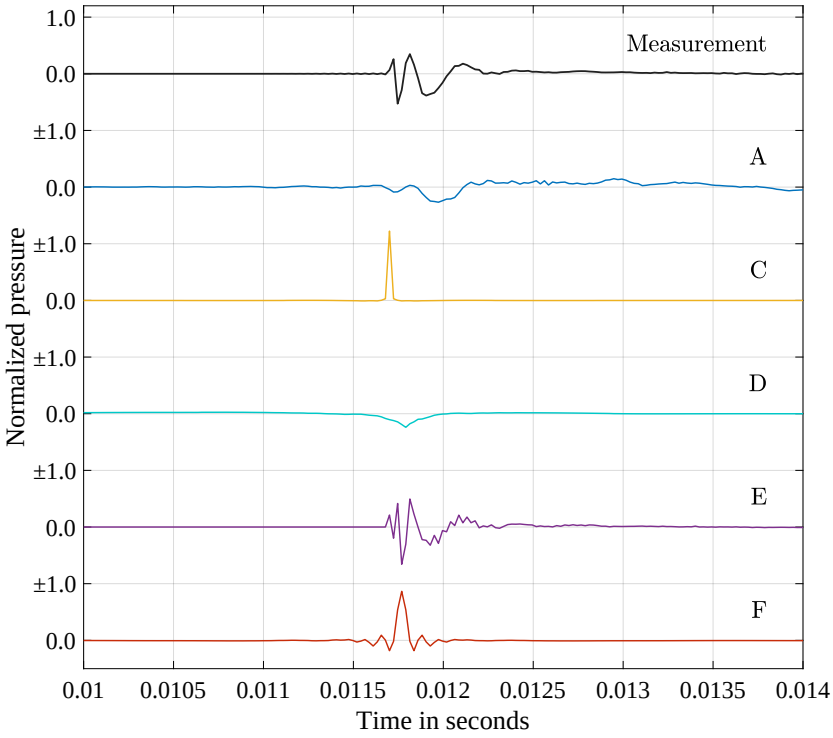


Figure B.4: Measured and simulated direct sound impulses of scene 3 for LS1–MP1. No bandpass filtering is applied.

Scene 5 (RS5): Diffraction (infinite wedge)

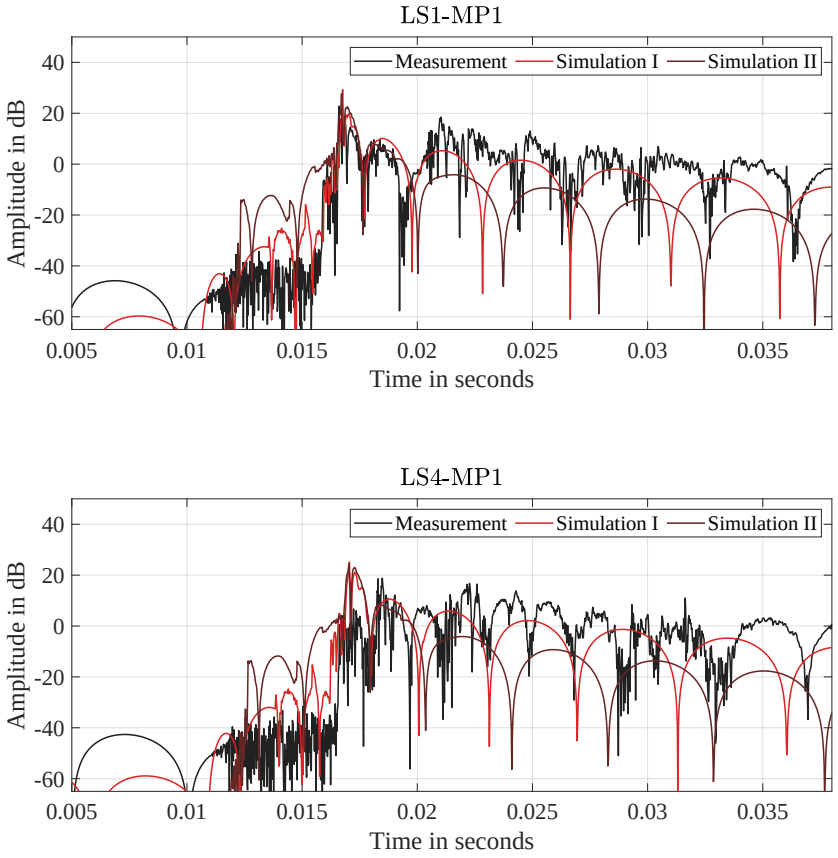


Figure B.5: Measured and simulated impulse responses of scene 5, diffraction on a partition

B.2 Complex rooms

Scene 9 (CR2): Small room (seminar room)

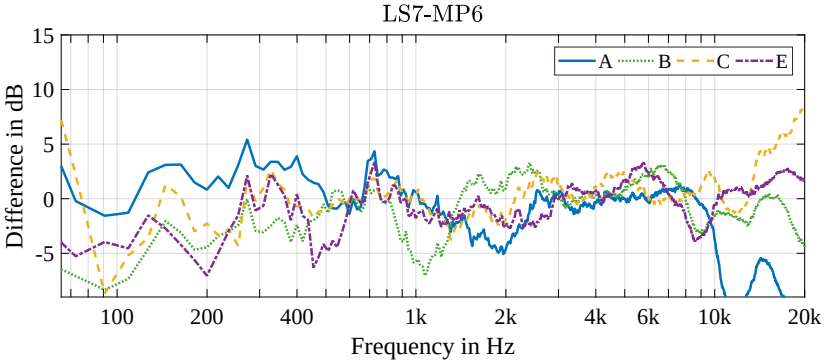


Figure B.6: Differences of measured and simulated spectrum of binaural room impulse responses (left channel only) for LS7–MP6. Only the early part up to 50 ms of the impulse responses was processed. A one-third octave smoothing filter was applied to both results before the spectral difference was calculated by means of division in the frequency domain.

Scene 10 (CR3): Medium room (chamber music hall)

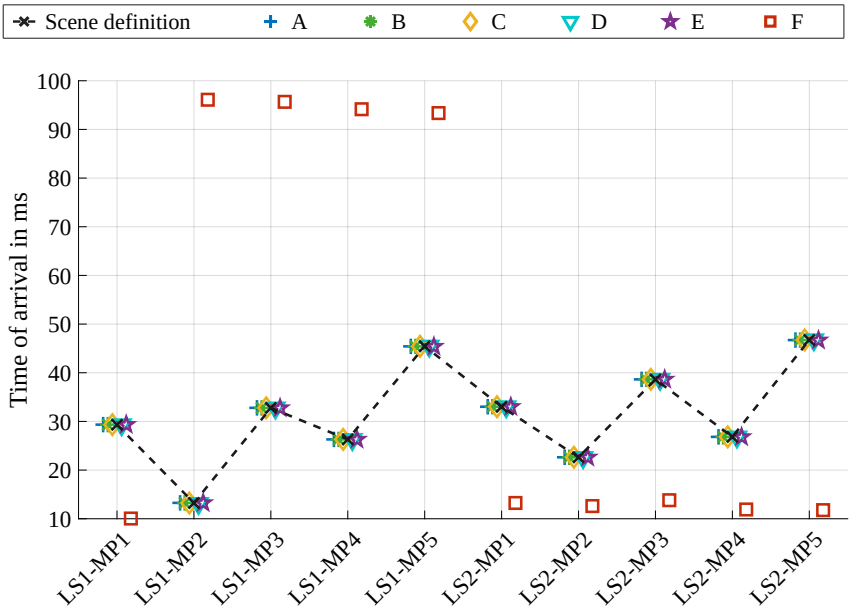


Figure B.7: Evaluated arrival times of all simulated room impulse responses of scene 10. For three participants (C, D and F), a constant time shift was applied as the submitted results included systematic delays.

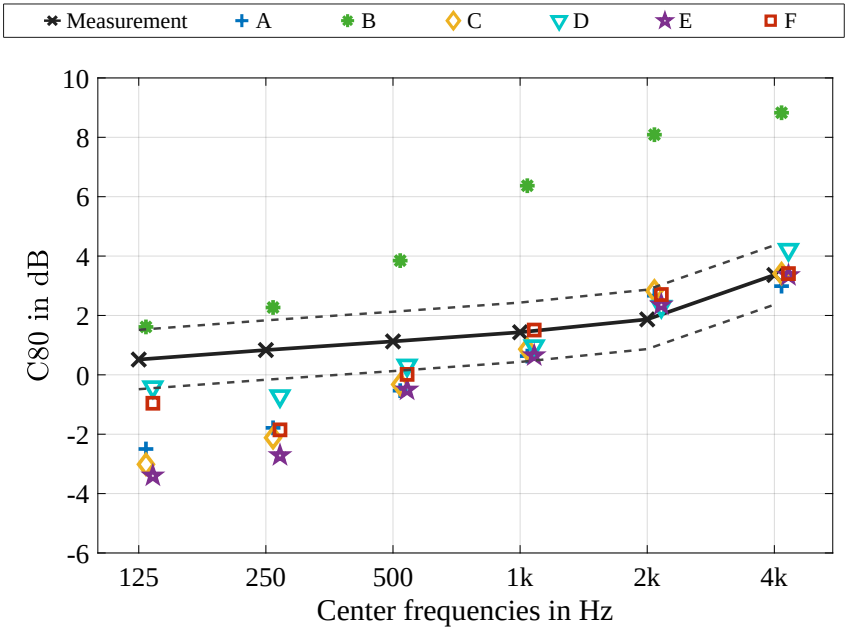


Figure B.8: Clarity parameter evaluated for scene 10, averaged for ten room impulse responses. Dashed lines indicate the just-noticeable difference of 1 dB.

Table B.1: Averaged early decay time values, calculated for six octave bands based on measured and simulated data of scene 10

	125 Hz	**250 Hz**	**500 Hz**	**1 kHz**	**2 kHz**	**4 kHz**
Measured	1.58 s	1.34 s	1.33 s	1.31 s	1.24 s	1.04 s
A	3.03 s	2.77 s	2.05 s	1.56 s	1.21 s	1.06 s
B	1.96 s	1.94 s	1.37 s	0.84 s	0.60 s	0.53 s
C	2.41 s	2.15 s	1.71 s	1.41 s	1.10 s	0.98 s
D	2.12 s	2.17 s	1.71 s	1.45 s	1.24 s	1.05 s
E	2.69 s	2.54 s	1.80 s	1.34 s	1.19 s	1.00 s
F	2.41 s	2.54 s	1.61 s	1.34 s	1.16 s	1.05 s

Table B.2: Averaged T20 values, calculated for six octave bands based on measured and simulated data of scene 10

	125 Hz	**250 Hz**	**500 Hz**	**1 kHz**	**2 kHz**	**4 kHz**
Measured	1.58 s	1.47 s	1.28 s	1.34 s	1.31 s	1.03 s
A	3.22 s	2.75 s	2.35 s	2.02 s	1.83 s	1.40 s
B	2.24 s	2.14 s	1.79 s	1.22 s	0.98 s	1.04 s
C	2.51 s	2.28 s	1.75 s	1.41 s	1.12 s	0.98 s
D	2.59 s	2.30 s	1.78 s	1.48 s	1.27 s	1.12 s
E	2.56 s	2.79 s	1.92 s	1.70 s	1.27 s	1.02 s
F	3.34 s	3.04 s	2.18 s	1.69 s	1.30 s	1.11 s

Scene 11 (CR4): Large room (auditorium)

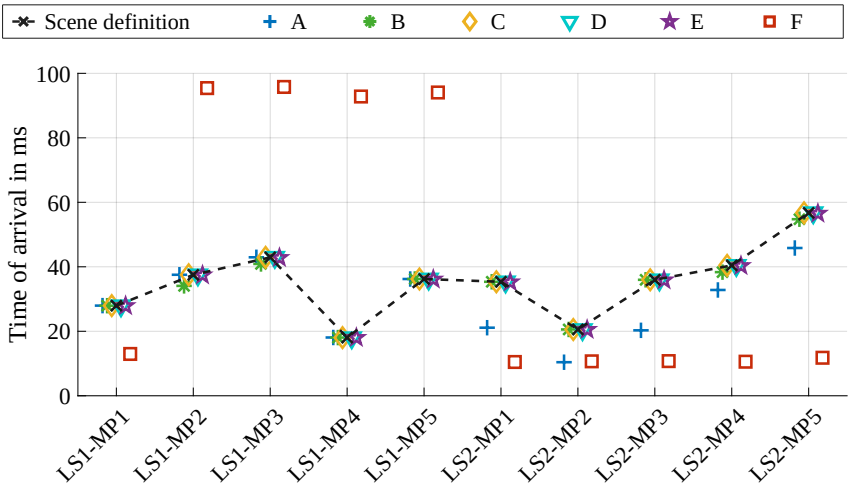


Figure B.9: Evaluated arrival times of all simulated room impulse responses of scene 11. For three participants (C, D and F), a constant time shift was applied as the submitted results included systematic delays.

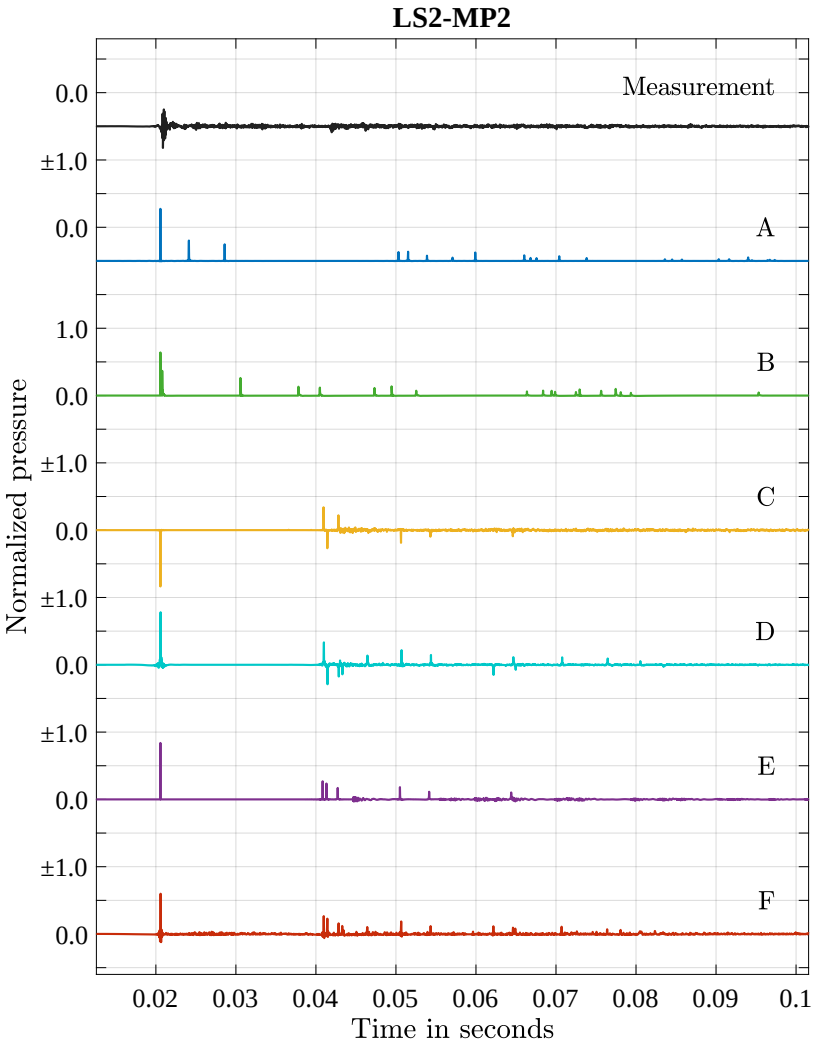


Figure B.10: Early part of measured and simulated impulse responses for LS2-MP2 of scene 11

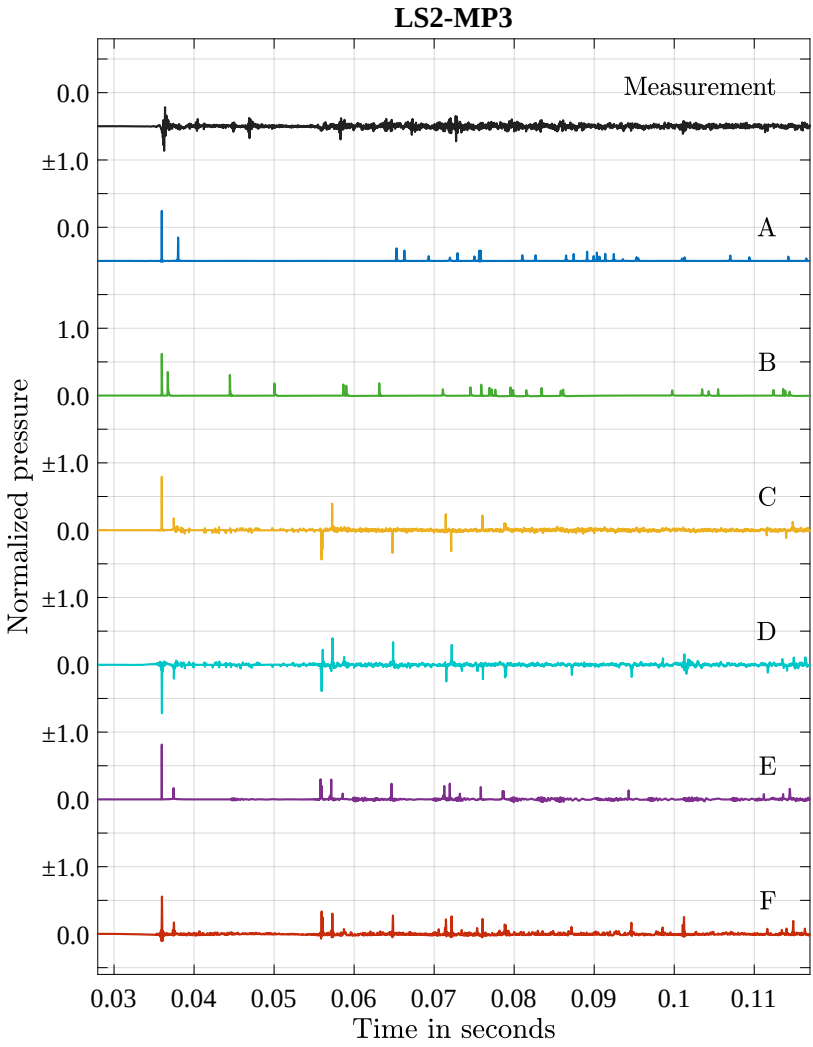


Figure B.11: Early part of measured and simulated impulse responses for LS2-MP3 of scene 11

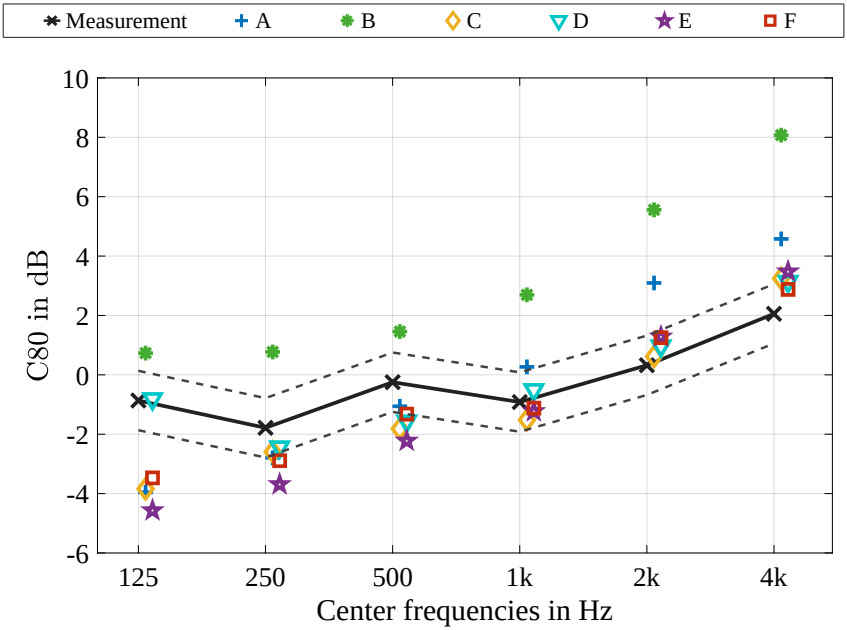


Figure B.12: Clarity parameter evaluated for scene 11, averaged for ten room impulse responses. Dashed lines indicate the just-noticeable difference of 1 dB.

C

Additional data related to informed simulations

C.1 Scene 1 (RS1): Single reflection (infinite plate)

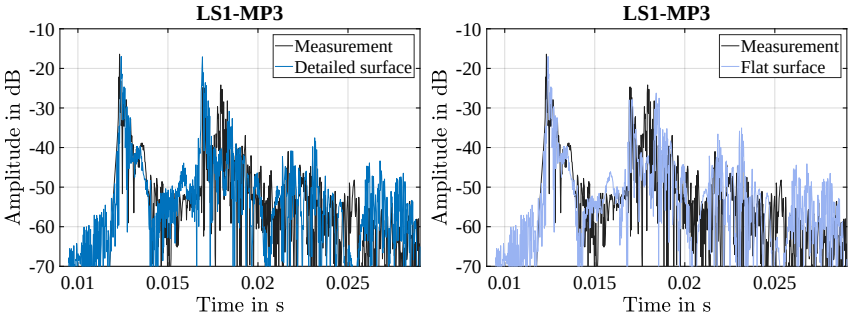


Figure C.1: Impulse responses for scene 1. Measured and simulated (detailed vs. flat surface) data for source–receiver combination LS1–MP3.

C.2 Scene 8 (CR1): Coupled rooms

Table C.1: Three input datasets used for scene 8, all data is applied for the 1 kHz octave frequency band. S denotes the surface area of the materials, A corresponds to the equivalent absorption area, α describes the absorption, and s the scattering coefficient of the listed materials.

Dataset I: Initial coefficients of BRAS database				
	S [m^2]	A [m^2]	α	s
painted concrete (R_2)	179.92	7.38	0.04	0.06
concrete (R_1)	150.41	5.41	0.04	0.02
tables & equipment (R_1)	35.95	9.03	0.25	0.57
absorber (R_1)	12.70	6.19	0.49	0.01

Dataset II: Fitted coefficients of BRAS database				
	S [m^2]	A [m^2]	α	s
painted concrete (R_2)	179.92	2.88	0.02	0.06
concrete (R_1)	150.41	5.41	0.04	0.02
tables & equipment (R_1)	35.95	9.02	0.25	0.57
absorber (R_1)	12.70	6.09	0.48	0.01

Dataset III: Iteratively matched to measured decay rates				
	S [m^2]	A [m^2]	α	s
painted concrete (R_2)	179.92	3.14	0.02	0.06
concrete (R_1)	150.41	13.14	0.09	0.31
tables & equipment (R_1)	35.95	6.95	0.19	0.57
absorber (R_1)	12.70	4.69	0.37	0.01

C.3 Reflection density of scene 9, 10 and 11

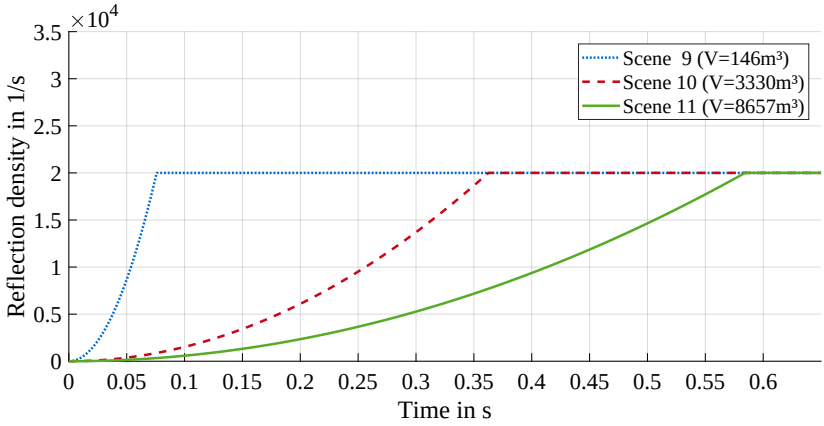


Figure C.2: Room reflection density for scene 9, 10 and 11 used for the filter synthesis of the simulated ray tracing result, calculated according to Eq. (5.7). The maximum reflection density is set to Nd_{max}=20000 s^{-1}.

C.4 Scene 9 (CR2): Small room (seminar room)

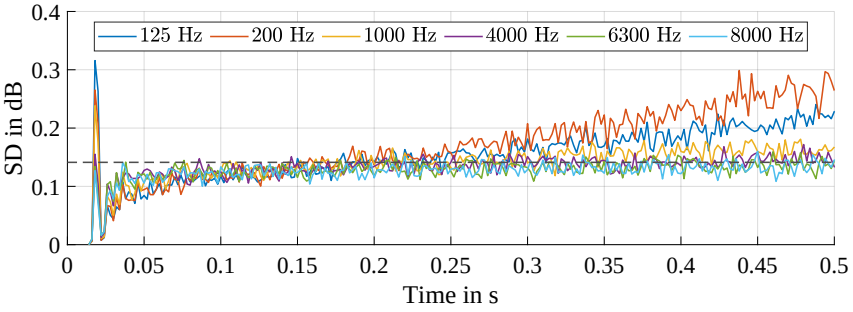


Figure C.3: Standard deviation (SD) of 100 simulated energy histograms for scene 9. Six one-third octave frequency bands are shown. Dashed grey line corresponds to the the calculated standard deviation (Eq. (5.6)), $\sigma_{\mathrm{EH}} = 0.14\,\mathrm{dB}$. Simulation configuration: $N{=}10^5$, $r_d{=}0.8\,\mathrm{m}$, $\Delta L_{max}{=}75\,\mathrm{dB}$, $\Delta t{=}2\,\mathrm{ms}$.

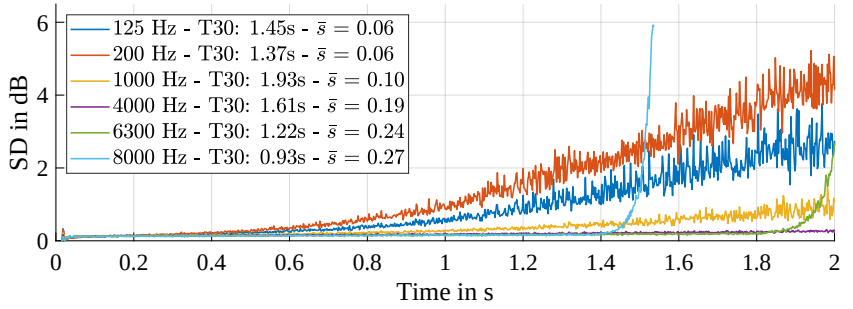


Figure C.4: Standard deviation (SD) of 100 simulated energy histograms for scene 9. Six one-third octave frequency bands are shown. Simulation configuration: $N{=}10^5$, $r_d{=}0.8\,\mathrm{m}$, $\Delta L_{max}{=}90\,\mathrm{dB}$, $\Delta t{=}2\,\mathrm{ms}$.

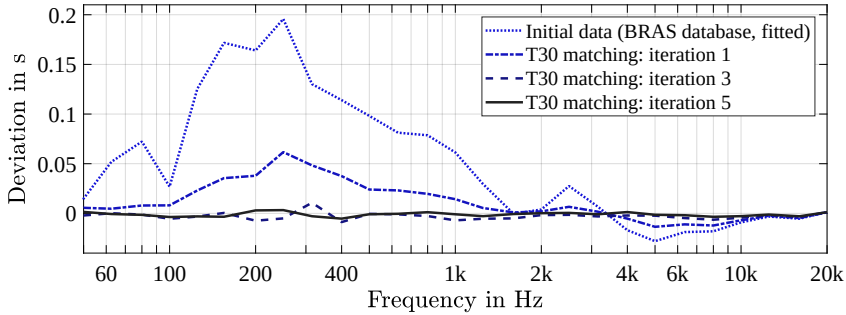


Figure C.5: Deviations from the measured values for simulated T30 values during the iterative matching process for scene 9.

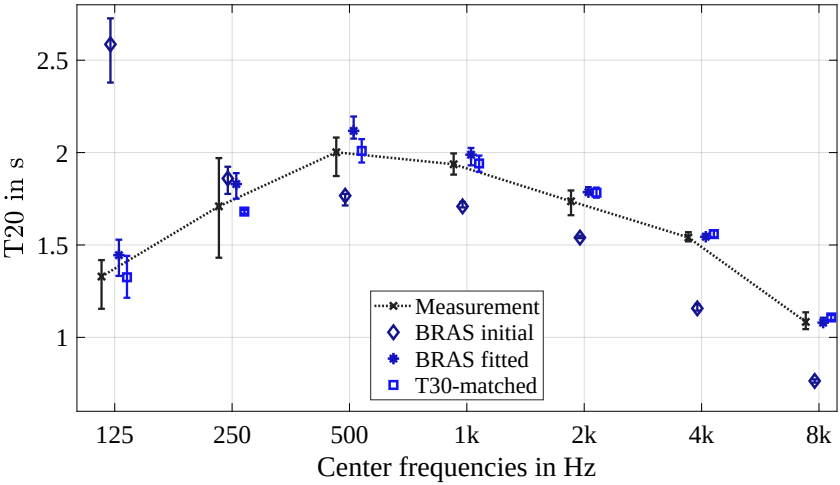


Figure C.6: Evaluated T20 values (average of 10 positions) for informed simulations and measurements of scene 9. Whiskers indicate the range from minimum to maximum value of the ten room impulse responses.

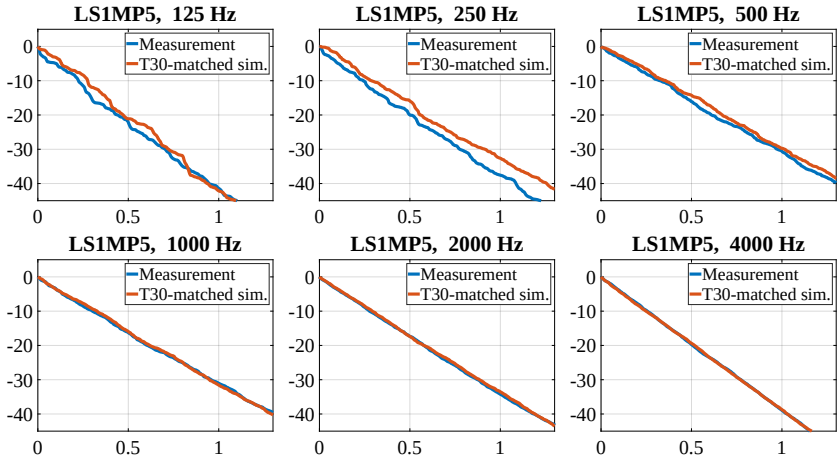


Figure C.7: Measured and simulated energy decay curves of scene 9, for LS1–MP5. The y-axis describes the energy in dB, x-axis the time in seconds.

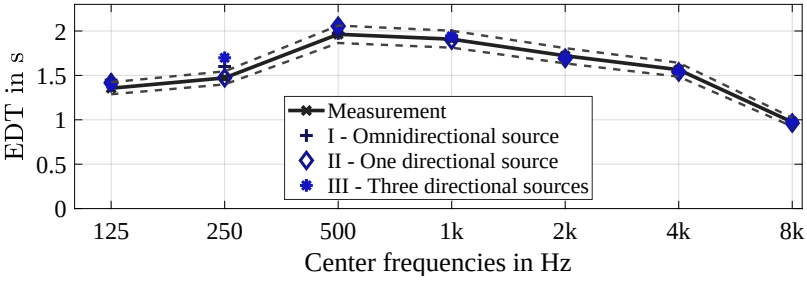


Figure C.8: Evaluated EDT values (average of 10 positions) for simulated and measured room impulse responses of scene 9. The sound sources in the simulations are configured in three different setups.

C.5 Scene 11 (CR4): Large room (auditorium)

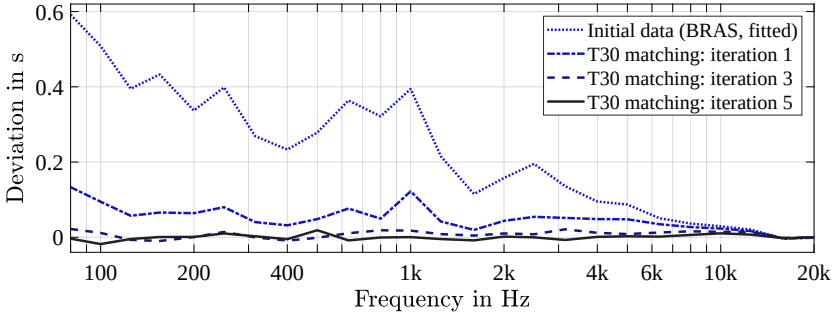


Figure C.9: Deviations from the measured values for simulated T30 values during the iterative matching process for scene 11

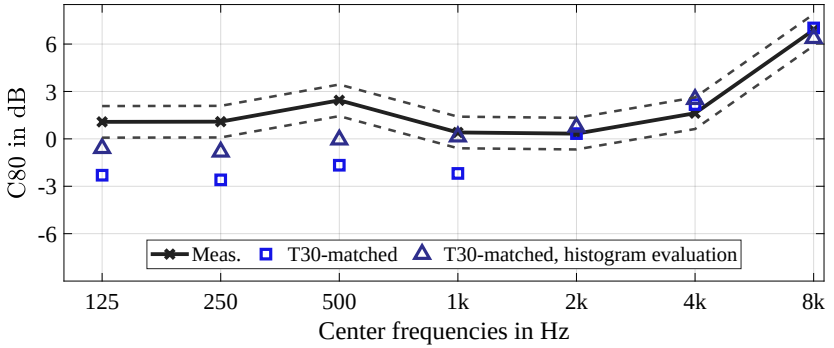


Figure C.10: Evaluated C80 values for simulated and measured room impulse responses of scene 11, positions LS2–MP2. Simulated values are either evaluated based on the impulse response or on the simulated energy histogram.

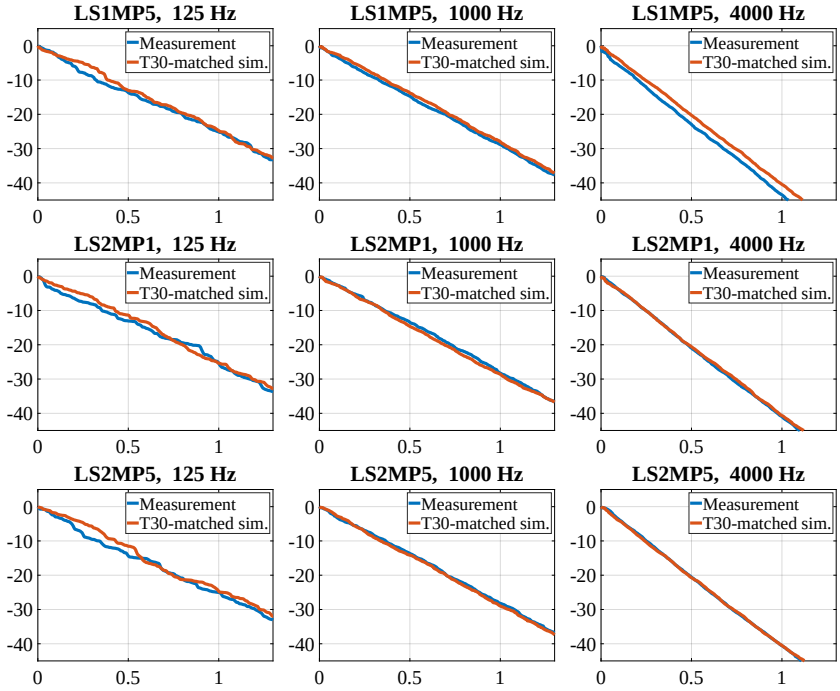


Figure C.11: Measured and simulated energy decay curves of scene 11. The y-axis describes the energy in dB, x-axis the time in seconds.

Bibliography

[1] Ahnert Feistel Media Group. EASE – User's Guide & Tutorial (Version 4.3), Acesssed in November 2018. http://www.afmg-support.de/SoftwareDownloadBase/AFMG/EASE/EASE_4.3_Tutorial_English.pdf (Accessed in November 2018).

[2] J. B. Allen and D. A. Berkley. Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Am.*, 65(4):943–950, Apr. 1979.

[3] H. Arau-Puchades. An improved reverberation formula. *Acta Acust. united Ac.*, 65(4):163–180, 1988. https://www.ingentaconnect.com/content/dav/aaua/1988/00000065/00000004/art00003.

[4] M. Aretz. *Combined wave and ray based room acoustic simulations of small rooms. Challenges and limitions on the way to realistic simulation results.* PhD thesis, RWTH Aachen University, Aachen, Germany, Sept. 2012. http://publications.rwth-aachen.de/record/197461/files/4317.pdf.

[5] M. Aretz, P. Dietrich, and M. Vorländer. Application of the mirror source method for low frequency sound prediction in rectangular rooms. *Acta Acust. united Ac.*, 100(2):306–319, Mar./Apr. 2014. doi:10.3813/AAA.918710.

[6] L. Aspöck, T. Maintz, and M. Vorländer. Determination of boundary conditions for room acoustics simulation by application of inverse calculation models. *J. Acoust. Soc. Am.*, 140(4):3128–3128, Nov 2016. http://publications.rwth-aachen.de/record/679281.

[7] L. Aspöck, S. Pelzer, and M. Vorländer. Using spatial information for the synthesis of the diffuse part of a binaural room impulse response. In *Fortschritte der Akustik – DAGA 2014: 40. Deutsche Jahrestagung für Akustik*, pages 71–72. Deutsche Gesellschaft für Akustik, 2014.

[8] L. Aspöck and M. Vorländer. Room geometry acquisition and processing methods for geometrical acoustics simulation models. In *9th Iberian Congress on Acoustics, EuroRegio 2016,*

Porto, Portugal, Jun 2016. https://pdfs.semanticscholar.org/40fa/3a99472519c2b9b83d09a66fa18867f0b276.pdf.

[9] L. Aspöck and M. Vorländer. Interaction of absorption and scattering coefficients in room simulations based on geometrical acoustics. In *Fortschritte der Akustik – DAGA 2019*, Mar 2019. http://publications.rwth-aachen.de/record/774351.

[10] L. Aspöck. Input and output data for informed room acoustic simulations of the BRAS scene database, 2020. doi:10.18154/RWTH-2020-09906.

[11] L. Aspöck, F. Brinkmann, D. Ackerman, S. Weinzierl, and M. Vorländer. GRAS – Ground Truth for Room Acoustical Simulation, Mar. 2018. doi:10.14279/depositonce-6726.

[12] L. Aspöck, F. Brinkmann, D. Ackerman, S. Weinzierl, and M. Vorländer. BRAS – Benchmark for Room Acoustical Simulation, May 2019. doi:10.14279/depositonce-6726.2.

[13] L. Aspöck, F. Brinkmann, D. Ackerman, S. Weinzierl, and M. Vorländer. BRAS – Benchmark for Room Acoustical Simulation, Oct. 2020. doi:10.14279/depositonce-6726.3.

[14] L. Aspöck, R. Opdam, and M. Vorländer. Acquisition of boundary conditions for a room acoustics simulation comparison. In *ISMRA*, Buenos Aires, Argentina, Sept. 11-13 2016. https://pdfs.semanticscholar.org/bf67/ac609b9d37212498c250a2b29cdae1f7d03b.pdf.

[15] L. Aspöck, S. Pelzer, F. Wefers, and M. Vorländer. A real-time auralization plugin for architectural design and education. In *Proc. of the EAA Joint Symposium on Auralization and Ambisonics*, pages 156–161, Berlin, Germany, Apr. 2014. doi:10.14279/depositonce-26.

[16] L. Aspöck and M. Vorländer. Synthesis of room impulse responses based on simulated energy decay curves. In *Fortschritte der Akustik – DAGA 2017*, pages 275–278, Kiel, Germany, Mar. 2017.

[17] L. Aspöck and M. Vorländer. Simulation of a coupled room scenario based on geometrical acoustics simulation models. *Proceedings of Meetings on Acoustics*, 36(1):015002, 2019. doi:10.1121/2.0001041.

[18] M. Barron. *Auditorium Acoustics and Architectural Design*. E & FN Spon, New York, NY, USA, 1993.

[19] H. E. Bass, L. C. Sutherland, A. J. Zuckerwar, D. T. Blackstock, and D. M. Hester. Atmospheric absorption of sound: Further developments. *J. Acoust. Soc. Am.*, 97(1):680–683, Jan. 1995. doi:10.1121/1.412989.

[20] G. Behler. An active loudspeaker point source for the measurement of high quality wide band room impulse responses. In *Proceedings of the Institute of Acoustics - Auditorium Acoustics 2018, Hamburg, Germany*, volume 40, pages 435–446, Hamburg, Germany, 2018. Institute of Acoustics.

[21] G. Benedetto and R. Spagnolo. Evaluation of sound absorbing coefficients in a reverberant room by computer-ray simulation. *Applied Acoustics*, 17(5):365 – 378, 1984. doi:10.1016/0003-682X(84)90048-3.

[22] G. Benedetto and R. Spagnolo. A study of barriers in enclosures by a ray-tracing computer model. *Applied Acoustics*, 17(3):183 – 199, 1984. doi:10.1016/0003-682X(84)90036-7.

[23] L. L. Beranek. *Concert halls and opera houses*. Springer, New York, USA, 2nd edition, 2004.

[24] B. Bernschütz. A spherical far field HRIR/HRTF compilation of the Neumann KU 100. In *AIA-DAGA 2013, International Conference on Acoustics*, pages 592–595, Merano, Italy, Mar. 2013.

[25] M. Berzborn, R. Bomhardt, J. Klein, J.-G. Richter, and M. Vorländer. The ITA-Toolbox: An open source MATLAB toolbox for acoustic measurements and signal processing. In *Fortschritte der Akustik – DAGA 2017*, pages 222–225, Kiel, Germany, Mar. 2017.

[26] M. A. Biot. Generalized boundary condition for multiple scatter in acoustic reflection. *J. Acoust. Soc. Am.*, 44(6):1616–1622, 1968. doi:10.1121/1.1911304.

[27] J. Blauert, H. Lehnert, J. Sahrhage, and H. Strauss. An interactive virtual-environment generator for psychoacoustic research. i: Architecture and implementation. *Acta Acust. united Ac.*, 86(1):94–102, 2000. https://www.ingentaconnect.com/contentone/dav/aaua/2000/00000086/00000001/art00014.

[28] J. Borish. Extension of the image model to arbitrary polyhedra. *J. Acoust. Soc. Am.*, 75(6):1827–1836, June 1984. doi:10.1121/1.390983.

[29] I. Bork. A comparison of room simulation software – The 2nd round robin on room acoustical computer simulation. *Acta Acust. united Ac.*, 86:943–956, 2000. https://www.ingentaconnect.com/content/dav/aaua/2000/00000086/00000006/art00008.

[30] I. Bork. Report on the 3rd round robin on room acoustical computer simulation - Part I: Measurements. *Acta Acust. united Ac.*, 91:740–752, 2005. https://www.ingentaconnect.com/contentone/dav/aaua/2005/00000091/00000004/art00015.

[31] I. Bork. Report on the 3rd round robin on room acoustical computer simulation - Part II: Calculations. *Acta Acust. united Ac.*, 91:753–763, 2005. https://www.ingentaconnect.com/contentone/dav/aaua/2005/00000091/00000004/art00016.

[32] M. A. Boucher. Phased geometrical acoustics using low / high frequency reflection coefficients with applications to absorption measurements, 2017.

[33] D. T. Bradley and L. M. Wang. Effect of model detail level on room acoustic computer simulations. *J. Acoust. Soc. Am.*, 111(5):2389–2389, 2002. doi:10.1121/1.4778118.

[34] D. T. Bradley and L. M. Wang. Quantifying the double slope effect in coupled volume room systems. *Building Acoustics*, 16(2):105–123, 2009. doi:10.1260/135101009788913275.

[35] J. S. Bradley. Some further investigations of the seat dip effect. *J. Acoust. Soc. Am.*, 90(1):324–333, 1991. doi:10.1121/1.401302.

[36] F. Brinkmann. *Binaural processing for the evaluation of acoustical environments*. Doctoral Thesis, Technical University Berlin, Berlin, Germany, 2019. doi:10.14279/depositonce-8510.

[37] F. Brinkmann, L. Aspöck, D. Ackermann, S. Lepa, M. Vorländer, and S. Weinzierl. A round robin on room acoustical simulation and auralization. *J. Acoust. Soc. Am.*, 145(4):2746–2760, 2019. doi:10.1121/1.5096178.

[38] F. Brinkmann, A. Lindau, M. Müller-Trapet, M. Vorländer, and S. Weinzierl. Cross-validation of measured and modeld head-related transfer functions. In *Fortschritte der Akustik – DAGA 2015*, pages 1118–1121, Nürnberg, Germany, Mar. 2015.

[39] F. Brinkmann, A. Lindau, and S. Weinzierl. On the authenticity of individual dynamic binaural synthesis. *J. Acoust. Soc. Am.*, 142(4):1784–1795, Oct. 2017. doi:10.1121/1.5005606.

[40] F. Brinkmann, A. Lindau, S. Weinzierl, G. Geissler, S. van de Par, M. Müller-Trapet, R. Opdam, and M. Vorländer. The FABIAN head-related transfer function data base. http://dx.doi.org/10.14279/depositonce-5718, Feb. 2017.

[41] F. Brinkmann, A. Lindau, S. Weinzierl, S. v. d. Par, M. Müller-Trapet, R. Opdam, and M. Vorländer. A high resolution and full-spherical head-related transfer function database for different head-above-torso orientations. *J. Audio Eng. Soc*, 65(10):841–848, Oct. 2017. doi:10.17743/jaes.2017.0033.

[42] Brüel & Kjær. Product Data Sheet - Concenser Microphone Cartridges – Types 4133 to 4181: https://www.bksv.com/media/doc/Bp0100.pdf (Accessed in November 2019).

[43] D. Cabrera, J. Xun, and M. Guski. Calculating reverberation time from impulse responses: A comparison of software implementations. *Acoustics Australia*, 44(2):369–378, 2016. doi:10.1007/s40857-016-0055-6.

[44] Y.-J. Choi and F. R. Fricke. A comparison of subjective assessments of recorded music and computer simulated auralizations in two auditoria. *Acta Acust. united Ac.*, 92(4):604–611, 2006. https://www.ingentaconnect.com/content/dav/aaua/2006/00000092/00000004/art00013.

[45] C. L. Christensen and J. H. Rindel. Estimating absorption of materials to match room model against existing room using a genetic algorithm. In *Forum Acusticum*, Kraków, Poland, Sept. 2014.

[46] T. J. Cox and P. D'Antonio. *Acoustic absorbers and diffusors. Theory, design and application.* Taylor & Francis, London, New York, 2nd edition, 2009.

[47] L. Cremer and H. A. Müller. *Die wissenschaftlichen Grundlagen der Raumakustik - Band 1.* Die wissenschaftlichen Grundlagen der Raumakustik. S. Hirzel Verlag Stuttgart, 2nd edition, 1978.

[48] B.-I. Dalenbäck. The importance of diffuse reflection in computerised room acoustic prediction and auralisation. In *Proc. of the Institute of Acoustics*, volume 17, pages 24–34. Institute of Acoustics, 1995.

[49] B.-I. Dalenbäck. Company Website presenting room simulation software CATT-Acoustic. http://www.catt.se (Accessed on 2020-03-23), Mar. 2020.

[50] B.-I. Dalenbäck, M. Kleiner, and P. Svensson. A macroscopic view of diffuse reflection. *J. Audio Eng. Soc*, 42(10):793–807, 1994. http://www.aes.org/e-lib/browse.cfm?elib=6927.

[51] J.-J. Embrechts. A geometrical acoustics approach linking surface scattering and reverberation in room acoustics. *Acta Acust. united Ac.*, 100(5):864–879, 2014. doi:doi:10.3813/AAA.918766.

[52] J. J. Embrechts and A. Billon. Theoretical determination of the random-incidence scattering coefficients of infinite rigid surfaces with a peridoic rectangular roughness profile. *Acta Acust. united Ac.*, 97(4):607–617, July 2011. doi:10.3813/AAA.918441.

[53] R. Falcon Perez, G. Götz, V. Pulkki, et al. Machine-learning-based estimation of reverberation time using room geometry for room effect rendering. In *Proceedings of the 23rd International Congress on Acoustics: 9-13 September 2019 in Aachen, Germany*, pages 7293–7300. Deutsche Gesellschaft für Akustik, Sep 2019. doi:10.18154/RWTH-CONV-239303.

[54] A. Farina. Simultaneous measurement of impulse response and distortion with swept-sine technique. In *108th AES Convention, Preprint 5093*, Paris, Feb. 2000. http://www.aes.org/e-lib/browse.cfm?elib=10211.

[55] A. Farina. Advancements in impulse response measurements by sine sweeps. In *Audio Engineering Society Convention 122*, May 2007. http://www.aes.org/e-lib/browse.cfm?elib=14106.

[56] D. Fitzroy. Reverberation formula which seems to be more accurate with nonuniform distribution of absorption. *J. Acoust. Soc. Am.*, 31(7):893–897, 1959. doi:10.1121/1.1907814.

[57] T. Funkhouser, N. Tsingos, I. Carlbom, G. Elko, M. Sondhi, J. E. West, G. Pingali, P. Min, and A. Ngan. A beam tracing method for interactive architectural acoustics. *J. Acoust. Soc. Am.*, 115(2):739–756, 2004. doi:10.1121/1.1641020.

[58] F. Georgiou, B. Briere de La Hosseraye, M. Hornikx, and P. Robinson. Design and simulation of a benchmark room for room acoustic auralizations. In *Proceedings of the 23rd International Congress on Acoustics: 9-13 September 2019 in Aachen, Germany*, pages 723–730. Deutsche Gesellschaft für Akustik, Sep 2019. doi:10.18154/RWTH-CONV-239684.

[59] G.R.A.S. Sound & Vibration. https://www.gras.dk/products/measurement-microphone-cartridge/product/ss_export/pdf2?product_id=163 (Accessed in November 2019).

[60] G. Grimm, J. Luberadzka, and V. Hohmann. Virtual acoustic environments for comprehensive evaluation of model-based hearing devices. *Int. J. Audiology*, 2016. doi:10.1080/14992027.2016.1247501.

[61] G. Grimm, J. Luberadzka, and V. Hohmann. A toolbox for rendering virtual acoustic environments in the context of audiology. *Acta Acust. united Ac.*, 105(3):566–578, 2019. doi:10.3813/AAA.919337.

[62] P. Guidorzi, L. Barbaresi, D. D'Orazio, and M. Garai. Impulse responses measured with mls or swept-sine signals applied to architectural acoustics: An in-depth analysis of the two methods and some case studies of measurements inside theaters. *Energy Procedia*, 78:1611 – 1616, 2015. doi:10.1016/j.egypro.2015.11.236. 6th International Building Physics Conference, IBPC 2015.

[63] R. E. Halliwell. Inter-laboratory variability of sound absorption measurement. *J. Acoust. Soc. Am.*, 73(3):880–886, 1983. doi:10.1121/1.389011.

[64] B. Hamilton. *Finite difference and finite volume methods for wave-based modelling of room acoustics*. PhD thesis, University of Edinburgh, Edinburgh, UK, 2016. http://www.brianhamilton.co/thesis/hamilton2016phdthesis.pdf.

[65] J. A. Hargreaves, L. R. Rendell, and Y. W. Lam. A framework for auralization of boundary element method simulations including source and receiver directivity. *J. Acoust. Soc. Am.*, 145(4):2625–2637, 2019. doi:10.1121/1.5096171.

[66] R. Heinz. Binaural room simulation based on an image source model with addition of statistical methods to include the diffuse sound scattering of walls and to predict the reverberant tail. *Applied Acoustics*, 38(2-4):145 – 159, 1993. doi:10.1016/0003-682X(93)90048-B.

[67] T. Hidaka and N. Nishihara. Objective evaluation of chamber-music halls in europe and japan. *J. Acoust. Soc. Am.*, 116(1):357–372, 2004. doi:10.1121/1.1760112.

[68] K. Hirosawa, K. Takashima, H. Nakagawa, M. Kon, A. Yamamoto, and W. Lauriks. Comparison of three measurement techniques for the normal

absorption coefficient of sound absorbing materials in the free field. *J. Acoust. Soc. Am.*, 126(6):3020–3027, Dec. 2009. doi:10.1121/1.3242355.

[69] M. Hodgson. Evidence of diffuse surface reflection in rooms. *J. Acoust. Soc. Am.*, 88(S1):S185–S185, 1990. doi:10.1121/1.2028828.

[70] M. Hodgson. On the accuracy of models for predicting sound propagation in fitted rooms. *J. Acoust. Soc. Am.*, 88(2):871–878, 1990. doi:10.1121/1.399737.

[71] M. Hornikx. Ten questions concerning computational urban acoustics. *Building and Environment*, 106:409 – 421, 2016. doi:10.1016/j.buildenv.2016.06.028.

[72] K. V. Horoshenkov, A. Khan, F.-X. Bécot, L. Jaouen, F. Sgard, A. Renault, N. Amirouche, F. Pompoli, N. Prodi, P. Bonfiglio, G. Pispola, F. Asdrubali, J. Hübelt, N. Atalla, C. K. Amédin, W. Lauriks, and L. Boeckx. Reproducibility experiments on measuring acoustical properties of rigid-frame porous media (round-robin tests). *J. Acoust. Soc. Am.*, 122(1):345–353, 2007. doi:10.1121/1.2739806.

[73] ISO 10534-2. *Determination of sound absorption coefficient and impedance in impedance tubes – Part 2: Transfer-function method.* International Organization for Standards, Geneva, Switzerland, Nov. 1998.

[74] ISO 17497-1. *Sound-scattering properties of surfaces. Part 1: Measurement of the random-incidence scattering coefficient in a reverberation room.* International Organization for Standards, Geneva, Switzerland, May 2004.

[75] ISO 17497-2. *Sound-scattering properties of surfaces. Part 2: Measurement of the directional diffusion coefficient in a free field.* International Organization for Standards, Geneva, Switzerland, May 2012.

[76] ISO 3382-1. *Measurement of room acoustic parameters – Part 1: Performance spaces.* International Organization for Standards, Geneva, Switzerland, 2009.

[77] ISO 354. *Measurement of sound absorption in a reverberation room.* International Organization for Standards, Geneva, Switzerland, May 2003.

[78] ISO/IEC Guide 98-3:2008. *Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement.* ISO, Geneva, Switzerland, 2008.

[79] K. Jambrosic, M. Horvat, and H. Domitrovic. Reverberation time measuring methods. In *Acoustics '08, 29 June - 4 July, 2008 in Paris, France*, pages 8511–8516, 2008. doi:10.1121/1.2934829.

[80] C.-H. Jeong, G. Marbjerg, and J. Brunskog. Uncertainty of input data for room acoustic simulations. In *Baltic-Nordic Acoustic Meeting (BNAM)*, Stockholm, Sweden, June 2016.

[81] B. F. G. Katz. International round robin on room acoustical impulse response analysis software. *Acoustic research letters online*, 5(4):158–164, Oct. 2004.

[82] M. Kleiner, B.-I. Dalenbäck, and P. Svensson. Auralization - An overview. *J. Audio Eng. Soc.*, 41(11):861–875, Nov. 1993.

[83] M. Kleiner, R. Orlowski, and J. Kirszenstein. A comparison between results from a physical scale model and a computer image source model for architectural acoustics. *Applied Acoustics*, 38(2):245 – 265, 1993. doi:10.1016/0003-682X(93)90055-B.

[84] C. Kosten. International comparison measurements in the reverberation room. *Acta Acust. united Ac.*, 10(5-6):400–411, 1960.

[85] A. Krokstad, S. Strom, and S. Sørsdal. Calculating the acoustical room response by the use of a ray tracing technique. *Journal of Sound and Vibration*, 8(1):118 – 125, 1968. doi:10.1016/0022-460X(68)90198-3.

[86] H. Kuttruff. Digital simulation of concert hall acoustics and its applications. *Acoustic Bulletin*, 16(5):5–8, 1991.

[87] H. Kuttruff. *Acoustics: An introduction*. Taylor & Francis, Milton Park, Abingdon, United Kingdom, 2007.

[88] H. Kuttruff. *Room acoustics*. Spon Press, Milton Park, Abingdon, United Kingdom, 5th edition edition, 2009.

[89] K. H. Kuttruff. Auralization of impulse responses modeled on the basis of ray-tracing results. *J. Audio Eng. Soc*, 41(11):876–880, 1993. http://www.aes.org/e-lib/browse.cfm?elib=6977.

[90] Y. W. Lam. A comparison of three diffuse reflection modeling methods used in room acoustics computer models. *J. Acoust. Soc. Am.*, 100(4):2181–2192, 1996. doi:10.1121/1.417927.

[91] A. Le Bot and A. Bocquillet. Comparison of an integral equation on energy and the ray-tracing technique in room acoustics. *J. Acoust. Soc. Am.*, 108(4):1732–1740, 2000. doi:10.1121/1.1287848.

[92] T. Lentz, D. Schröder, M. Vorländer, and I. Assenmacher. Virtual reality system with integrated sound field simulation and reproduction. *EURASIP Journal of Advances in Signal Processing*, pages 1–19, 2007.

[93] A. Lindau. *Binaural resynthesis of acoustical environments. Technology and perceptual evaluation.* Ph.D. thesis, Technical Universtiy Berlin, Berlin, Germany, May 2014.

[94] A. Lindau, V. Erbes, S. Lepa, H.-J. Maempel, F. Brinkmann, and S. Weinzierl. A Spatial Audio Quality Inventory (SAQI). *Acta Acust. united Ac.*, 100(5):984–994, 2014.

[95] A. Lindau, L. Kosanke, and S. Weinzierl. Perceptual evaluation of model- and signal-based predictors of the mixing time in binaural room impulse responses. *J. Audio Eng. Soc.*, 60(11):887–898, Nov. 2012.

[96] A. Lindau and S. Weinzierl. Assessing the plausibility of virtual acoustic environments. *Acta Acust. united Ac.*, 98(5):804–810, Sept./Oct. 2012. doi:10.3813/AAA.918562.

[97] J. Llorca, I. B. Witew, E. Redondo, and M. Vorländer. 3d Modelling Photogrammetry to Support Acoustic Measurements and Derive Geometries for Simulation. In *Proceedings of the Institute of Acoustics*, pages 661–666. Auditorium Acoustics 2018, Hamburg (Germany), 4 Oct 2018 - 6 Oct 2018, Institute of Acoustics, Oct 2018.

[98] P. Luizard and B. F. G. Katz. Coupled volume multi-slope room impulse responses: A quantitative analysis method. In *Intl. Conf. on Auditorium Acoustics*, volume 33, pages 169–176, Dublin, Ireland, 2011. https://hal.archives-ouvertes.fr/hal-01789731.

[99] P. Luizard and B. F. G. Katz. Investigation of the effective aperture area of sliding and hinged doors between coupled spaces. *J. Acoust. Soc. Am.*, 136(2):EL135–EL141, Aug. 2014. doi:10.1121/1.4890202.

[100] P. Luizard, M. Otani, J. Botts, L. Savioja, and B. F. G. Katz. Comparison of sound field measurements and predictions in coupled volumes between numerical methods and scale model measurements. In *Proceedings of Meetings on Acoustics*, volume 19, Montreal, Canada, June 2013. doi:10.1121/1.4799138.

[101] G. Marbjerg, J. Brunskog, C.-H. Jeong, and E. Nilsson. Development and validation of a combined phased acoustical radiosity and image source model for predicting sound fields in rooms. *J. Acoust. Soc. Am.*, 138(3):1457–1468, 2015. doi:10.1121/1.4928297.

[102] R. S. Martín, M. Arana, J. Machín, and A. Arregui. Impulse source versus dodecahedral loudspeaker for measuring parameters derived from the impulse response in room acoustics. *J. Acoust. Soc. Am.*, 134(1):275–284, July 2013. doi:10.1121/1.4808332.

[103] S. Müller and P. Massarani. Transfer-function measurement with sweeps. *J. Audio Eng. Soc*, 49(6):443–471, 2001. http://www.aes.org/e-lib/browse.cfm?elib=10189.

[104] E. Mommertz. Angle-dependent in-situ measurements of reflection coefficients using a subtraction technique. *Applied Acoustics*, 46(3):251 – 263, 1995. doi:10.1016/0003-682X(95)00027-7. Building Acoustics.

[105] B. Mondet, J. Brunskog, C.-H. Jeong, and J. H. Rindel. From absorption to impedance: Enhancing boundary conditions in room acoustic simulations. *Applied Acoustics*, 157:106884, 2020. doi:10.1016/j.apacoust.2019.04.034.

[106] M. Müller-Trapet. *Measurement of surface reflection properties*. PhD thesis, RWTH Aachen University, Aachen, Germany, July 2015.

[107] M. Müller-Trapet, P. Dietrich, M. Aretz, J. van Gemmeren, and M. Vorländer. On the in situ impedance measurement with pu-probes — Simulation of the measurement setup. *J. Acoust. Soc. Am.*, 134(2):1082–1089, Aug. 2013. doi:10.1121/1.4812250.

[108] R. V. Muñoz, L. Aspöck, and J. Fels. Spatial release from masking under different reverberant conditions in young and elderly subjects: Effect of moving or stationary maskers at circular and radial conditions. *Journal of Speech, Language, and Hearing Research*, 62(9):3582–3595, 2019. doi:10.1044/2019_JSLHR-H-19-0092.

[109] J. M. Navarro, J. J. Lopez, and J. Escolano. Estimation of absorption coefficients values of surface materials using a diffusion equation model. In *Proceedings of Meetings on Acoustics*, volume 19, page 015101, 2013. doi:10.1121/1.4798984.

[110] G. Naylor and J. H. Rindel. Predicting room acoustical behavior with the odeon computer model. *J. Acoust. Soc. Am.*, 92(4):2346–2346, 1992. doi:10.1121/1.404931.

[111] R. Neubauer and B. Kostek. Prediction of the reverberation time in rectangular rooms with non-uniformly distributed sound absorption. *Archives of Acoustics*, 26(3), 2001.

[112] C. Nocke. In-situ acoustic impedance measurement using a free-field transfer function method. *Applied Acoustics*, 59(3):253 – 264, Mar. 2000. doi:10.1016/S0003-682X(99)00004-3.

[113] M. Noisternig, T. Carpentier, and O. Warusfel. Espro 2.0 – implementation of a surrounding 350-loudspeaker array for 3d sound field reproduction. In *Audio Engineering Society Conference: UK 25th Conference: Spatial Audio in Today's 3D World*, Mar 2012. http://www.aes.org/e-lib/browse.cfm?elib=18121.

[114] E.-M. Nosal, M. Hodgson, and I. Ashdown. Improved algorithms and methods for room sound-field prediction by acoustical radiosity in arbitrary polyhedral rooms. *J. Acoust. Soc. Am.*, 116(2):970–980, 2004. doi:10.1121/1.1772400.

[115] Odeon A/S. ODEON Room Acoustics Software User's Manual. https://odeon.dk/wp-content/uploads/2017/09/ODEON_Manual.pdf (Accessed in May 2019), Sept. 2017.

[116] J. Ohm and H. Lüke. *Signalübertragung: Grundlagen der digitalen und analogen Nachrichtenübertragungssysteme*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2010.

[117] A. M. Ondet and J. L. Barbry. Modeling of sound propagation in fitted workshops using ray tracing. *J. Acoust. Soc. Am.*, 85(2):787–796, 1989. doi:10.1121/1.397551.

[118] A. V. Oppenheim, R. W. Schafer, and J. R. Buck. *Discrete-time signal processing*. Prentice Hall, Upper Saddle, USA, 2nd edition edition, 1999.

[119] T. Otsuru, T. Sakuma, and S. Sakamoto. Constructing a database of computational methods for environmental acoustics. *Acoustical Sci. & Tech.*, 26(2):221–224, 2005. doi:10.1250/ast.26.221.

[120] F. Pausch, L. Aspöck, M. Vorländer, and J. Fels. An extended binaural real-time auralization system with an interface to research hearing aids for experiments on subjects with hearing loss. *Trends in Hearing*, 22, 2018. doi:10.1177/2331216518800871.

[121] A. Pedrero, M. Ángeles Navacerrada, D. de la Prida, L. Iglesias, and A. Díaz-Chyla. On the accuracy of the sound absorption measurement with an impedance gun. *Applied Acoustics*, 158:107039, 2020. doi:10.1016/j.apacoust.2019.107039.

[122] S. Pelzer, M. Aretz, and M. Vorländer. Quality assessment of room acoustic simulation tools by comparing binaural measurements and simulations in an optimized test scenario. In *Proceedings of Forum Acusticum 2011 : 27 June - 01 July, Aalborg, Denmark*. European Acoustics Association (EAA), 2011. https://pdfs.semanticscholar.org/9759/9e27f93c340fd2caf0c76204e4dedf4439b5.pdf.

[123] S. Pelzer, L. Aspöck, D. Schröder, and M. Vorländer. Integrating real-time room acoustics simulation into a cad modeling software to enhance the architectural design process. *Building Acoustics*, 4(2):113–138, Apr. 2014. doi:10.3390/buildings4020113.

[124] S. Pelzer, B. Sanches Masiero, and M. Vorländer. 3D Reproduction of Room Auralizations by Combining Intensity Panning, Crosstalk Cancellation and Ambisonics. In *Proceedings of the EAA Joint Symposium on Auralization and Ambisonics 2014 / Hrsg.: Stefan Weinzierl ; Michael Vorländer ; Franz Zotter ; Hans-Joachim Maempel ; Alexander Lindau*, pages 182–188, Berlin, 2014. Universitätsbibliothek Technische Universität Berlin. doi:10.14279/depositonce-33.

[125] S. Pelzer and M. Vorländer. Inversion of a room acoustics model for the determination of acoustical surface properties in enclosed spaces. In *Proc. Mtgs. Acoust.*, volume 19, page 015115, June 2013. doi:10.1121/1.4800297.

[126] S. Pelzer, M. Vorländer, and H.-J. Maempel. Room modeling for acoustic simulation and auralization tasks: Resolution of structural detail. In *Fortschritte der Akustik – DAGA 2010*, pages 15–18, Berlin, Germany, mar 2010.

[127] S. Pelzer and M. Vorländer. Frequency- and Time-dependent Geometry for Real-time Auralizations. In *Proceedings ICA 2010, 20th International Congress on Acoustics : 23 - 27 August 2010, Sydney, New South Wales, Australia*. Australian Acoustical Society, NSW Division, 2010. https://publications.rwth-aachen.de/record/128719.

[128] Physikalisch-Technische Bundesanstalt (PTB). Tabularized absortion and scattering values. (Accessed on 2019-12-14), 2019.

https://www.ptb.de/cms/fileadmin/internet/fachabteilungen/abteilung_1/1.6_schall/1.63/abstab_wf.zip.

[129] J. Picaut and N. FORTIN. I-Simpa, a graphical user interface devoted to host 3D sound propagation numerical codes. In S. F. d'Acoustique, editor, *Acoustics 2012*, Nantes, France, Apr. 2012. https://hal.archives-ouvertes.fr/hal-00810893.

[130] A. Pilch. Optimization in the validation of the room acoustic model. In *9th Iberian Congress on Acoustics, EuroRegio 2016, Porto, Portugal*, Jun 2016.

[131] F. Pompoli, P. Bonfiglio, K. Horoshenkov, A. Khan, L. Jaouen, F.-X. Becot, F. Sgard, F. Asdrubali, F. D' Alessandro, J. Hübelt, N. Atalla, C. Amédin, W. Lauriks, and L. Boeckx. A round robin test of the acoustical properties of porous media. *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, 255(6):1702–1706, 2017. https://www.ingentaconnect.com/content/ince/incecp/2017/00000255/00000006/art00084.

[132] B. N. J. Postma and B. F. G. Katz. Creation and calibration method of acoustical models for historic virtual reality auralizations. *Virtual Reality*, 19(3):161–180, Nov 2015. doi:10.1007/s10055-015-0275-3.

[133] B. N. J. Postma and B. F. G. Katz. Perceptive and objective evaluation of calibrated room acoustic simulation auralizations. *J. Acoust. Soc. Am.*, 140(6):4326–4337, Dec. 2016. doi:10.1121/1.4971422.

[134] W. Probst. Determination of absorption coefficients in a "virtual" reverberation chamber – an application of acoustic simulation with the sound particle model. *Lärmbekämpfung*, 10(5):204 – 211, 2015.

[135] B. Rafaely. *Fundamentals of spherical array processing*. Springer, Berlin, Heidelberg, Germany, 1st edition, 2015. doi:10.1007/978-3-662-45664-4.

[136] D. D. Rife and J. Vanderkooy. Transfer-function measurement with maximum-length sequences. *J. Audio Eng. Soc*, 37(6):419–444, 1989. http://www.aes.org/e-lib/browse.cfm?elib=6086.

[137] J. Rindel. Acoustic design of reflectors in auditoria. In *Proc. Inst. Acoust.(UK)*, volume 14 (2), pages 119–128, 1992.

[138] J. H. Rindel. Attenuation of sound reflections due to diffraction. In *Proceedings of Nordic Acoustical Meeting*, pages 257–260, Aalborg, Denmark, 1986.

[139] J. H. Rindel. Room acoustic modelling techniques: A comparison of a scale model and a computer model for a new opera theatre. *Building Acoustics*, 18(3-4):259–280, 2011. doi:10.1260/1351-010X.18.3-4.259.

[140] J. H. Rindel, G. B. Nielsen, and C. L. Christensen. Diffraction around corners and over wide barriers in room acoustic simulations. In *Proceedings of the 16th International Congress on Sound Vibration*, Jul 2009.

[141] T. Rossing. *Springer Handbook of Acoustics*. Springer Handbook of Acoustics. Springer New York, 2007. doi:10.1007/978-0-387-30425-0.

[142] W. C. Sabine. *Collected papers on acoustics*. Books on Architecture. Havard University Press, Cambridge, 1922.

[143] T. Sakuma, S. Sakamoto, and T. Otsuru. A benchmarking framework for wave-based computational methods in architectural acoustics. In *Proc. Proc. 18th Intl. Congress on Acoustics (ICA)*, pages 281–284, 2004. https://www.icacommission.org/Proceedings/ICA2004Kyoto/pdf/Mo4.B2.2.pdf.

[144] T. Sakuma, S. Sakamoto, and T. Otsuru. *Computational simulation in architectural and environmental acoustics. Methods and applications of wave-based computation*. Springer, Tokyo et al., 2014. doi:10.1007/978-4-431-54454-8.

[145] T. Sakuma, P. Svensson, A. Franck, and S. Sakamoto. A round-robin test program on wave-based computational methods for room-acoustic analysis. In *Proc. Forum Acusticum, Sevilla*. Citeseer, 2002.

[146] R. San Martín, A. Arregui, J. Machín, and M. Arana. Comparison of measured and simulated room acoustic parameter values using high resolution grids. *Acoustics Australia*, 42(1), 2014.

[147] L. Savioja, J. Hiipakka, and T. Lokki. A framework for evaluating virtual acoustic environments. In *Audio Engineering Society Convention 110*, May 2001. http://www.aes.org/e-lib/browse.cfm?elib=9987.

[148] L. Savioja and U. P. Svensson. Overview of geometrical room acoustic modeling techniques. *J. Acoust. Soc. Am.*, 138(2):708–730, Aug. 2015. doi:10.1121/1.4926438.

[149] C. Schissler, A. Nicholls, and R. Mehra. Efficient HRTF-based spatial audio for area and volumetric sources. *IEEE Transactions on Visualization and Computer Graphics*, 22(4):1356–1366, Apr. 2016. doi:10.1109/TVCG.2016.2518134.

[150] D. Schröder and T. Lentz. Real-time processing of image sources using binary space partitioning. *J. Audio Eng. Soc*, 54(7/8):604–619, 2006. http://www.aes.org/e-lib/browse.cfm?elib=13899.

[151] D. Schröder. *Physically based real-time auralization of interactive virtual environments*. PhD thesis, RWTH Aachen, Aachen, Germany, 2011. http://publications.rwth-aachen.de/record/50580/files/3875.pdf.

[152] D. Schröder and A. Pohl. Real-time hybrid simulation method including edge diffraction. In *EAA Auralization Symposium 2009: Espoo, Finland, 15th-17th June*, 2009.

[153] D. Schröder and M. Vorländer. Raven: A real-time framework for the auralization of interactive virtual environments. In *Forum Acusticum*, pages 1541–1546. Aalborg Denmark, 2011. https://www2.ak.tu-berlin.de/~akgroup/ak_pub/seacen/2011/Schroeder_2011b_P2_RAVEN_A_Real_Time_Framework.pdf.

[154] M. R. Schroeder. New method of measuring reverberation time. *J. Acoust. Soc. Am.*, 37:409–412, 1965. doi:10.1121/1.1939454.

[155] M. R. Schroeder. Digital simulation of sound transmission in reverberant spaces. *J. Acoust. Soc. Am.*, 47(2A):424–431, 1970. doi:10.1121/1.1911541.

[156] M. R. Schroeder and K. H. Kuttruff. On frequency response curves in rooms. comparison of experimental, theoretical, and monte carlo results for the average frequency spacing between maxima. *J. Acoust. Soc. Am.*, 34(1):76–80, 1962. doi:10.1121/1.1909022.

[157] T. J. Schultz and B. G. Watters. Propagation of sound across audience seating. *J. Acoust. Soc. Am.*, 36(5):885–896, May 1964. doi:10.1121/1.1919111.

[158] SEACEN Research Unit. Project website: International round robin on room simulation and auralization. http://rr.auralisation.net (last accessed on 2019-05-23). http://rr.auralisation.net/.

[159] K. Sekiguchi and T. Hanyu. Study on acoustic index variations due to small changes in the observation point. In *15th International Congress on Acoustics*, pages 2121–2122, Seattle, USA, 1998.

[160] M. Skålevik. Low frequency limits of reflector arrays. In *6th International Conference on Auditorium Acoustics 2006, Copenhagen, Denmark*, volume 28, pages 223–230. Institute of Acoustics, 2006. https://akutek.info/Papers/MS_Array_2007.pdf.

[161] M. C. Soares, F. O. Taminato, T. d. P. Assis, T. A. O. Roxo, P. G. Lima, and R. A. Tenenbaum. Estudo comparativo entre medições e simulações de respostas impulsivas biauriculares de salas para distintas rotações da cabeça sobre o torso. In *In Proceedings of the XXVIII Encontro da Sociedade Brasileira de Acústica (SOBRAC) in Porto Alegre, Brasil*, Campinas, Brasil, Oct. 2018. doi:10.17648/sobrac-87153.

[162] A. Southern, S. Siltanen, D. T. Murphy, and L. Savioja. Room impulse response synthesis and validation using a hybrid acoustic model. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(9):1940–1952, Sept. 2013. doi:10.1109/TASL.2013.2263139.

[163] F. Spandöck. Akustische modellversuche. *Annalen der Physik*, 412(4):345–360, 1934. doi:10.1002/andp.19344120402.

[164] U. M. Stephenson. An energetic approach for the simulation of diffraction within ray tracing based on the uncertainty relation. *Acta Acust. united Ac.*, 96(3):516–535, 2010. doi:10.3813/AAA.918304.

[165] P. Svensson and U. R. Kristiansen. Computational modelling and simulation of acoutic spaces. In *AES 22nd International Conference: Virtual, Synthetic, and Entertainment Audio*, Espoo, Finland, June 2002.

[166] R. A. Tenenbaum, F. O. Taminato, and V. S. G. Melo. Room acoustics modeling using a hybrid method with fast auralization with artificial neural network techniques. In *Proceedings of the 23rd International Congress on Acoustics: 9-13 September 2019 in Aachen, Germany*. Deutsche Gesellschaft für Akustik, Sep 2019. http://pub.dega-akustik.de/ICA2019/data/articles/000002.pdf.

[167] R. A. Tenenbaum, F. O. Taminato, and V. S. G. Melo. Fast auralization using radial basis functions type of artificial neural network techniques. *Applied Acoustics*, 157:106993, 2020. doi:10.1016/j.apacoust.2019.07.041.

[168] E. Tijs. *Study and development of an in situ acoustic absorption measurement method*. PhD thesis, University of Twente, Enschede, The Netherlands, Apr. 2013.

[169] F. C. Tommasini, O. A. Ramos, M. X. Hüg, and S. P. Ferreyra. A computational model to implement binaural synthesis in a hard real-time auditory virtual environment. *Acoustics Australia*, 47(1):51–66, Apr 2019. doi:10.1007/s40857-019-00152-7.

[170] J. C. B. Torres, L. Aspöck, and M. Vorländer. Comparative study of two geometrical acoustic simulation models. *J. Brazilian Soc. of Mechanical Sciences and Engineering*, 40(6):300, 5 2018. doi:10.1007/s40430-018-1226-1.

[171] R. R. Torres, U. P. Svensson, and M. Kleiner. Computation of edge diffraction for more accurate room acoustics auralization. *J. Acoust. Soc. Am.*, 109(2):600–610, Feb. 2001. doi:10.1121/1.1340647.

[172] N. Tsingos, I. Carlbom, G. Elko, R. Kubli, and T. Funkhouser. Validating acoustical simulations in the Bell Labs Box. *IEEE Computer Graphics and Applications*, 22(4):28–37, July 2002. doi:10.1109/MCG.2002.1016696.

[173] M. Vercammen. Improving the accuracy of sound absorption measurement according to ISO 354. In *Proc. Int. Symp. on Room Acoustics (ISRA)*, Melbourne, Australia, 2010. http://www.acoustics.asn.au/conference_proceedings/ICA2010/cdrom-ISRA2010/Papers/P2d.pdf.

[174] M. Vorländer. Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm. *J. Acoust. Soc. Am.*, 86(1):172–178, July 1989. doi:10.1121/1.398336.

[175] M. Vorländer. *Untersuchungen zur Leistungsfähigkeit des raumakustischen Schallteilchenmodells*. PhD thesis, RWTH Aachen University, Aachen, Germany, Feb. 1989.

[176] M. Vorländer. International round robin on room acoustical computer simulations. In *15th International Congress on Acoustics*, pages 689–692, Trondheim, Norway, June 1995.

[177] M. Vorländer. *Auralization. Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality*. Springer, Berlin, Heidelberg, Germany, 1st edition, 2008. doi:10.1007/978-3-540-48830-9.

[178] M. Vorländer. Computer simulations in room acoustics: Concepts and uncertainties. *J. Acoust. Soc. Am.*, 133(3):1203–1213, Mar. 2013. doi:10.1121/1.4788978.

[179] H. Wang, I. Sihar, R. Pagán Muñoz, and M. Hornikx. Room acoustics modelling in the time-domain with the nodal discontinuous galerkin method. *J. Acoust. Soc. Am.*, 145(4):2650–2663, 2019. doi:10.1121/1.5096154.

[180] L. M. Wang, J. Rathsam, and S. Ryherd. Interactions of model detail level and scattering coefficients in room acoustic computer simulation. In *Intl. Symposium on Room Acoustics, April 11-13, 2004 in Hyogo, Japan*, 2004. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1017&context=archengfacpub.

[181] A. Weber and B. F. G. Katz. Numerical simulation round robin of a coupled volume case: Preliminary results. In *Proceedings of the 23rd International Congress on Acoustics: 9-13 September 2019 in Aachen, Germany*. Deutsche Gesellschaft für Akustik, Sep 2019. https://hal.archives-ouvertes.fr/hal-02304623/file/001446.pdf.

[182] F. Wefers and D. Schröder. Real-time auralization of coupled rooms. In *EAA Auralization Symposium 2009; Espoo, Finland, 15th-17th June*, 2009.

[183] S. Weinzierl. *Handbuch der Audiotechnik*. Springer, Heidelberg, Germany, 1st Edition edition, 2008. doi:10.1007/978-3-540-34301-1.

[184] S. Weinzierl, P. Sanvito, F. Schultz, and C. Büttner. The acoustics of Renaissance theatres in Italy. *Acta Acust. united Ac.*, 101(3):632–641, May 2015. doi:10.3813/AAA.918858.

[185] T. Wendt, S. van de Par, and S. D. Ewert. A computationally-efficient and perceptually-plausible algorithm for binaural room impulse response simulation. *J. Audio Eng. Soc*, 62(11):748 – 766, Nov. 2014. doi:10.17743/jaes.2014.0042.

[186] T. Wendt, S. van de Par, and S. D. Ewert. Perceptually plausible acoustics simulation of single and coupled rooms. *J. Acoust. Soc. Am.*, 140(4):3178–3178, 2016. doi:10.1121/1.4969989.

[187] I. B. Witew, P. Dietrich, D. de Vries, and M. Vorländer. Uncertainty of room acoustics measurements - How many measurement positions are necessary to describe the conditions in auditoria? In *Proc. Int. Symp. on Room Acoustics (ISRA)*, Melbourne, Australia, Aug. 2010.

[188] I. B. Witew and M. Vorländer. The uncertainty of room impulse response measurements. In *Proceedings of the 23rd International Congress on Acoustics: 9-13 September 2019 in Aachen, Germany*. Deutsche Gesellschaft für Akustik, Sep 2019.

[189] N. Xiang, Y. Jing, and A. C. Bockman. Investigation of acoustically coupled enclosures using a diffusion-equation model. *J. Acoust. Soc. Am.*, 126(3):1187–1198, 2009. doi:10.1121/1.3168507.

[190] M. Yadav, D. Cabrera, W. Martens, et al. Auditory room size perceived from a room acoustic simulation with autophonic stimuli. *Acoustics Australia*, 39(3), 2011. http://www.acoustics.asn.au/journal/2011/2011_39_3_Yadav.pdf.

Acknowledgments

Various factors and helpful persons have supported and motivated me to write this thesis. The field of acoustics was very interesting to me when I took my first lectures in this area in 2008, and it still is today. The main reason for this is Prof. Michael Vorländer: Thank you, not only for supervising my thesis, but also for constantly increasing my interest in acoustics-related topics and for the many discussions, challenges and opportunities.

In 2009, I started my first student job at ITA: Thanks to Dirk Schröder for opening the doors into the world of room acoustic simulation. Being part of the virtual acoustics group at ITA always meant a very good working atmosphere for me. I constantly learned new things and shared great experiences with my colleagues, especially with Sönke Pelzer, Frank Wefers, Jonas Stienen and Michael Kohnen. Thanks also to Gottfried Behler, Karin Charlier, Janina Fels and Rolf Kaldenbach for your support in all sorts of matters and for making the institute to the special place it is, along with all fellow colleagues and students at ITA.

The BRAS project and the related round robin would not exist without Fabian Brinkmann and David Ackermann of the Audio Communication Group at TU Berlin. Big thanks to you for the efficient cooperation and the friendly atmosphere at all times. Our job got easier as we also had help of many others: Rob Opdam, Thomas Maintz, Philipp Eschbach, Ingo Witew, Johannes Klein, Marco Berzborn, Oliver Strauch, Hannes Helmholtz and Omid Kokabi. Special thanks to the workshop of ITA, lead by Uwe Schlömer, for the precise and reliable work.

I am also very thankful to Stefan Weinzierl for co-supervising my thesis, but also for a lot of valuable feedback and for leading the DFG research unit SEACEN, which mostly funded my position at ITA.

Finally, I would like to thank my parents, who paved the way for me to start studying at a university, and, most of all my wife, for your love, your patience and for always having my back, especially in the final phase of the thesis.

Bisher erschienene Bände der Reihe

Aachener Beiträge zur Akustik

ISSN 1866-3052
ISSN 2512-6008 (seit Band 28)

1	Malte Kob	Physical Modeling of the Singing Voice ISBN 978-3-89722-997-6 40.50 EUR
2	Martin Klemenz	Die Geräuschqualität bei der Anfahrt elektrischer Schienenfahrzeuge ISBN 978-3-8325-0852-4 40.50 EUR
3	Rainer Thaden	Auralisation in Building Acoustics ISBN 978-3-8325-0895-1 40.50 EUR
4	Michael Makarski	Tools for the Professional Development of Horn Loudspeakers ISBN 978-3-8325-1280-4 40.50 EUR
5	Janina Fels	From Children to Adults: How Binaural Cues and Ear Canal Impedances Grow ISBN 978-3-8325-1855-4 40.50 EUR
6	Tobias Lentz	Binaural Technology for Virtual Reality ISBN 978-3-8325-1935-3 40.50 EUR
7	Christoph Kling	Investigations into damping in building acoustics by use of downscaled models ISBN 978-3-8325-1985-8 37.00 EUR
8	Joao Henrique Diniz Guimaraes	Modelling the dynamic interactions of rolling bearings ISBN 978-3-8325-2010-6 36.50 EUR
9	Andreas Franck	Finite-Elemente-Methoden, Lösungsalgorithmen und Werkzeuge für die akustische Simulationstechnik ISBN 978-3-8325-2313-8 35.50 EUR

10	Sebastian Fingerhuth	Tonalness and consonance of technical sounds ISBN 978-3-8325-2536-1 42.00 EUR
11	Dirk Schröder	Physically Based Real-Time Auralization of Interactive Virtual Environments ISBN 978-3-8325-2458-6 35.00 EUR
12	Marc Aretz	Combined Wave And Ray Based Room Acoustic Simulations Of Small Rooms ISBN 978-3-8325-3242-0 37.00 EUR
13	Bruno Sanches Masiero	Individualized Binaural Technology. Measurement, Equalization and Subjective Evaluation ISBN 978-3-8325-3274-1 36.50 EUR
14	Roman Scharrer	Acoustic Field Analysis in Small Microphone Arrays ISBN 978-3-8325-3453-0 35.00 EUR
15	Matthias Lievens	Structure-borne Sound Sources in Buildings ISBN 978-3-8325-3464-6 33.00 EUR
16	Pascal Dietrich	Uncertainties in Acoustical Transfer Functions. Modeling, Measurement and Derivation of Parameters for Airborne and Structure-borne Sound ISBN 978-3-8325-3551-3 37.00 EUR
17	Elena Shabalina	The Propagation of Low Frequency Sound through an Audience ISBN 978-3-8325-3608-4 37.50 EUR
18	Xun Wang	Model Based Signal Enhancement for Impulse Response Measurement ISBN 978-3-8325-3630-5 34.50 EUR
19	Stefan Feistel	Modeling the Radiation of Modern Sound Reinforcement Systems in High Resolution ISBN 978-3-8325-3710-4 37.00 EUR
20	Frank Wefers	Partitioned convolution algorithms for real-time auralization ISBN 978-3-8325-3943-6 44.50 EUR

21	Renzo Vitale	Perceptual Aspects Of Sound Scattering In Concert Halls ISBN 978-3-8325-3992-4 34.50 EUR
22	Martin Pollow	Directivity Patterns for Room Acoustical Measurements and Simulations ISBN 978-3-8325-4090-6 41.00 EUR
23	Markus Müller-Trapet	Measurement of Surface Reflection Properties. Concepts and Uncertainties ISBN 978-3-8325-4120-0 41.00 EUR
24	Martin Guski	Influences of external error sources on measurements of room acoustic parameters ISBN 978-3-8325-4146-0 46.00 EUR
25	Clemens Nau	Beamforming in modalen Schallfeldern von Fahrzeuginnenräumen ISBN 978-3-8325-4370-9 47.50 EUR
26	Samira Mohamady	Uncertainties of Transfer Path Analysis and Sound Design for Electrical Drives ISBN 978-3-8325-4431-7 45.00 EUR
27	Bernd Philippen	Transfer path analysis based on in-situ measurements for automotive applications ISBN 978-3-8325-4435-5 50.50 EUR
28	Ramona Bomhardt	Anthropometric Individualization of Head-Related Transfer Functions Analysis and Modeling ISBN 978-3-8325-4543-7 35.00 EUR
29	Fanyu Meng	Modeling of Moving Sound Sources Based on Array Measurements ISBN 978-3-8325-4759-2 44.00 EUR
30	Jan-Gerrit Richter	Fast Measurement of Individual Head-Related Transfer Functions ISBN 978-3-8325-4906-0 45.50 EUR
31	Rhoddy A. Viveros Muñoz	Speech perception in complex acoustic environments: Evaluating moving maskers using virtual acoustics ISBN 978-3-8325-4963-3 35.50 EUR
32	Spyros Brezas	Investigation on the dissemination of unit watt in airborne sound and applications ISBN 978-3-8325-4971-8 47.50 EUR

33	Josefa Oberem	Examining auditory selective attention: From dichotic towards realistic environments ISBN 978-3-8325-5101-8 43.00 EUR
34	Johannes Klein	Directional Room Impulse Response Measurement ISBN 978-3-8325-5139-1 41.00 EUR
35	Lukas Aspöck	Validation of room acoustic simulation models ISBN 978-3-8325-5204-6 57.00 EUR