LUND UNIVERSITY

Room Acoustic Ray-Tracing

Understanding, evaluating and expanding the toolset

Autio, Hanna

2024

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

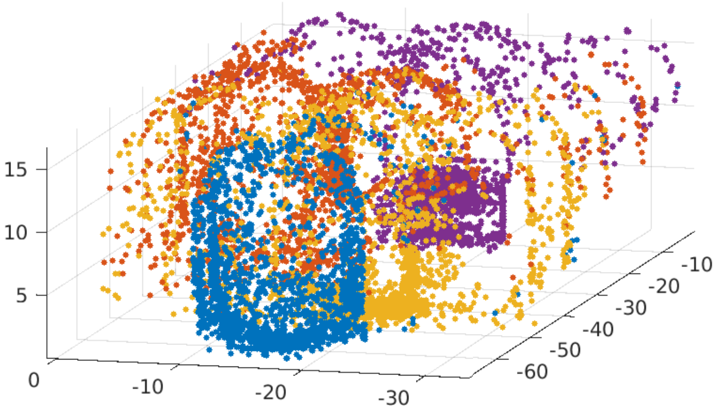
Citation for published version (APA):
Autio, H. (2024). *Room Acoustic Ray-Tracing: Understanding, evaluating and expanding the toolset*. Department of Construction Sciences, Lund University.

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

LUND
UNIVERSITY



ROOM ACOUSTIC RAY-TRACING
Understanding, evaluating and expanding the toolset

HANNA AUTIO

Engineering
Acoustics

Doctoral Thesis

DEPARTMENT OF CONSTRUCTION SCIENCES
DIVISION OF ENGINEERING ACOUSTICS
ISRN LUTVDG/TVBA--24/1018--SE (1-174) | ISSN 0281-8477
ISBN 978-91-8039-902-9 (print) | ISBN 978-91-8039-903-6 (pdf)
DOCTORAL THESIS

ROOM ACOUSTIC RAY-TRACING

Understanding, evaluating and expanding the toolset

HANNA AUTIO

Printed by V-husets tryckeri LTH, Lund, Sweden, February 2024 *(Pl).*

For information, address:
Div. of Engineering Acoustics, Faculty of Engineering LTH,
Lund University, Box 118, SE-221 00 Lund, Sweden.
Homepage: http://www.akustik.lth.se

Acknowledgements

The thesis work described in this paper has been performed at the Division of Engineering Acoustics and the Division of Structural mechanics at the Department of Construction Sciences at LTH. The work has been funded mainly by VR grant 2016-01784, *Multisensoriska perspektiv på det senmedeltida Vadstena kloster*.

As with any project, my work on this thesis has been a journey. This thesis is the product of a lot of work and thought, and although I will not be able to mention everyone that has helped me through this, a few of you must be mentioned.

First of all, thank you to the Vadstena team, Eva Lindqvist Sandgren, Karin Strinnholm Lagergren, Carolina Ask, Stefan Lindgren and Delphine Bard, for bringing me on. Thank you for trusting me in my role as a scientist even before I knew anything about what I was doing, and for letting me be a part of the project. I would also like to thank the greater community that has helped us with the historic reconstruction work, for engaging with us and showing me that my work mattered to someone.

Thank you to all my supervisors, Delphine Bard, Eva Lindqvist Sandgren, Susanne Heyden, Mathias Barbagallo and Erling Nilsson. A special thanks to Susanne Heyden and Eva Lindqvist Sandgren for sticking with me. There would not have been a thesis without you.

I would also like to thank all my co-authors, particularly Mathias and Nikos. I am deeply grateful that you made time for me and my work, and cooperating with you has been a key part in moving my research forward. Thank you also to Emma Arvidsson, Juan Negreira and Erik Molin.

The division of structural mechanics has also been a red thread through my time working on this thesis. Thank you, Erik Serrano, for giving me the chance to finish this work. I am also deeply grateful to Karin Forsman, for being a voice of reason in a sometimes nonsensical academic world.

Finally, thank you to my family and friends. Thank you Hampus, my partner, for entertaining my weirdly specific questions on how to best plot some really particular data and for keeping me fed. Thank you also to Alve, my son, for reminding me of the pure joy in exploring and trying to make sense of the world. Thank you also to my parents Imma and Jari, my siblings Linda and Toffe and my niece Natalie and nephew Sebastian. Your love and support keeps me safe and sane.

And finally, I would like to dedicate this work to Varga and Flinta.

Abstract

Room acoustic simulations are an important tool, both for design of spaces with satisfactory acoustics and for use in virtual reality experiences. Auralization of acoustic simulations are becoming more and more important in the context of Virtual and Augmented reality (VR and AR respectively), as the technology is used more and more in healthcare, culture and research. In this work, ray tracing simulations for room acoustics were studied. New methods were introduced to improve their overall performance in terms of simulation speed and accuracy for the purpose of auralization in VR.

Acoustic ray tracing simulations and auralizations of the historically important Vadstena abbey church were performed. The heritage site was reconstructed digitally based on historical sources. Analysis of the simulation results showed that the sound field was characterized by acoustic subspaces where the activities of the monastical congregations were better supported than in the nave. Redistribution of acoustic energy by the vaulted ceiling was found to have a significant impact on the transmission of sound into the nave. Accordingly, the distribution and redirection of acoustic energy was identified as important factors in the simulation. Scattering plays a central role in those processes and was consequently studied further in the project. It was additionally found that faster simulations were necessary to achieve good performance for VR.

A room acoustic ray tracer was implemented to facilitate the development and testing of new algorithms. The ray tracer runs on the GPU and can generate an estimate of the energetic room impulse response.

Using the developed ray tracer, some typical algorithms for scattering were tested and evaluated based on their simulation accuracy, simulation speed and usability. It was found that randomly determining whether rays are scattered or geometrically reflected was the overall best method. This is also the most commonly used algorithm in ray tracers today. An extension to this model is suggested, so that so-called 1D-scatterers can be modelled more accurately. The extended algorithm was tested by simulation of a room where the orientation of 1D-scatterers on the walls had been determined to lead to audible variations in room acoustic parameters. The algorithm for directional scattering was capable of emulating those variations.

Finally, a strategy to maintain simulation accuracy under strong restrictions on computational resources was presented. The method is called Iterative ray tracing, and uses a fundamental property of Monte Carlo simulations methods to improve simulation precision over time. By combining the results of multiple ray tracing passes, the overall number of rays (or samples) is increased, and thus variance is decreased while the expected outcome is maintained.

In all, this research has shed light on the historical acoustics of Vadstena abbey church, developed mathematical models for the analysis of ray tracing simulations, and produced new strategies for use in room acoustic ray tracing simulations.

Populärvetenskaplig sammanfattning

Ljud omger människan hela tiden, från tiden innan hon fötts tills dödsögonblicket. För hörande individer är ljud ett viktigt verktyg för att förstå och tolka omgivningen, och en tillfredsställande ljudmiljö är viktig för det allmänna välbefinnandet. Simulering av ljud och auralisering, alltså att generera hörbart ljud utifrån digitala beräkningar, är viktiga verktyg för att skapa goda ljudmiljöer, både när nya byggnader ska utformas och i virtuella upplevelser.

Det finns många sätt att simulera ljud, men särskilt intressant i samband med mänskliga ljudupplevelser är rumsakustiska simuleringar. En av de vanligaste metoderna för rumsakustiska simuleringar är så kallade strålgångs-beräkningar, eller ray tracing. Rumsakustiska strålgångssimuleringar syftar till att efterlikna ljudmiljön inne i ett rum, baserat på ljudkällor, lyssnare och rummets utseende. Resultaten kan användas till auraliseringar, till exempel som ett verktyg inom Virtual Reality (VR) eller Augmented Reality (AR). I VR bär användaren ett headset som avskiljer användaren från verkligheten och visar både bild och ljud ifrån en virtuell omgivning, medan AR-teknik istället används för att utöka verkligheten med digitala element på till exempel genomskinliga glasögon.

Auralisering kan också användas för att ta del av kulturella upplevelser och besöka historiska platser. En del av arbetet som beskrivs i den här avhandlingen har varit att genomföra ljudsimulering med auraliseringar för en historisk rekonstruktion av Vadstena klosterkyrka i Sverige. Klosterkyrkan, som länge representerade en av de mest inflytelserika makthavarna i regionen, studerades av en tvärvetenskaplig grupp forskare inom projektet "Multisensoriska perspektiv på det senmedeltida Vadstena kloster". Detta arbete resulterade i en digital modell av klosterkyrkan som användes för ljudsimuleringar, och det historiska ljudfältet studerades.

Utifrån simuleringsresultaten noterades det att klosterkyrkans ljudfält påverkades mycket av så kallad *acoustic scattering* (ungefär akustisk spridning). Inom akustiken används termen för att beskriva när ljudvågor splittras och sprids ut i rummet, särskilt vid reflektion mot en ojämn yta. I Vadstena klosterkyrka orsakades scattering särskilt av takvalven och av ett stort antal religiösa föremål i mittskeppet. Med detta som utgångspunkt har simuleringsmetoder för scattering varit ett fokusområde i det här arbetet.

En gedigen undersökning av olika metoder för att simulera scattering genomfördes. Denna gav vid handen att den vanligaste metoden också var mest lämpad ur ett helhetsperspektiv, eftersom den var relativt snabb, förutsägbar och korrekt. I ett senare steg utvidgades den algoritmen så att fler typer av ytor kan modelleras, specifikt ytor som reflekterar ljud mer i en viss riktning. Exempel på en sådana ytor skulle vara kyrkbänkar, tak med ribbvalv eller trappor. För att testa metoden användes mätningar där det uppmätts skillnad i ljudupplevelsen beroende på om en sådan yta var orienterad vertikalt eller horisontellt. Algoritmen kunde efterlikna de uppmätta skillnaderna.

En annan aspekt av akustiksimulering och auralisering för VR är beräkningstid. Inom VR måste uppdateringsfrekvensen vara hög, och akustiska simuleringar tar ofta för lång tid för att fungera fullt ut i VR idag. Inom det här projektet har en metod utvecklats som kallats Iterative ray tracing, (ungefär Iterativ strålgångs-beräkning). Metoden kan användas för att se till att den totala simuleringskvaliteten inte påverkas av begränsningar i tillgängligt minne eller beräkningstid.

Inom ramarna för det här projektet har också flera statistiska modeller utvecklats, både för mätresultat och för simuleringsresultat. Detta, sammantaget med en ordentlig utvärdering av olika algoritmer för scattering, har bidragit till en fördjupad förståelse för strålgångsmodellen i allmänhet.

Sammantaget har arbetet i avhandlingen producerat de första auraliseringarna av 1400-talets Vadstena klosterkyrka. Dessutom har nya statistiska modeller utvecklats, som ger nya och utökade analytiska verktyg för att beskriva resultaten av ljudsimuleringar och mätningar. Till sist har två nya algoritmer som kan användas inom rumsakustisk simulering utvecklats, en för att möjliggöra förenklad simulering av ytor som sprider ljud särskilt i en viss riktning, och en för att bibehålla simuleringens kvalitet när beräkningsresurserna begränsas.

Popular summary

Sounds surround humans constantly, from the time before birth to the moment of death. For hearing individuals, sound is an important tool for understanding and interpreting their surroundings, and a satisfactory sound environment is important for their overall comfort. Sound simulations and auralization, i.e. generating audible sound from digital data, are important tools to create good sound environments both when designing new buildings and in virtual experiences.

There are many reasons to simulate sound, but particularly relevant to human sound experiences are room acoustic simulations. One of the most common methods for room acoustic simulations are ray tracing simulations. Room acoustic ray tracing aims at emulating the sound field within a room based on sound sources, listeners and the room configuration. The simulation results can be used for auralizations, for example in virtual reality (VR).

Auralization can also be used to partake in cultural experiences and visit historical places. Part of the work described in this dissertation has been to perform acoustic simulations and auralizations for a historic reconstruction of Vadstena abbey church in Sweden. The church, which acted as a hub for political and cultural influence during the late Middle Ages, was studied by a cross-disciplinary research team within a project titled “Multisensory perspectives on Vadstena abbey in the late Middle Ages”. The work resulted in a digital model of the abbey church, which was used for room acoustic ray tracing simulations and subsequent analysis of the historical sound field.

Based on the results of the room acoustic simulation, it was noted that the sound field in the church was affected by acoustic scattering. The term acoustic scattering refers to the process when sound waves are split up and spread out into a room, particularly after reflection from a rough surface. In Vadstena abbey church, scattering was caused primarily by the vaulted ceiling and the large number of religious objects in the nave. Building on from these results, simulation algorithms for scattering has been an area of focus in this work.

Some common algorithms for scattering were tested and evaluated thoroughly. The research showed that the most common method also was the best method overall, as it was relatively fast, predictable and accurate. At a later stage, this method was expanded to enable simulation of other types of surfaces, particularly surfaces which scatter sound primarily in one dimension. Examples of such surfaces are church benches, battened ceilings or stairs. The new method was tested by comparison to measurements where differences could be seen based on whether such surfaces were oriented horizontally or vertically. The algorithm was capable of emulating those differences.

If acoustic simulations should be used in VR, the calculation time is an important factor. In VR, the update rate must be high, and acoustic simulations typically take too long to be useful in VR today. Within this project, a strategy called Iterative ray tracing has been developed. The strategy can be used to ensure that the overall simulation accuracy is not affected by restrictions in terms of calculation time or digital memory.

Within this project, several mathematical-statistical models have been developed, both for simulations results and for measurements. Combined with a thorough review of different scattering algorithms, this has contributed to a deepened understanding of the room acoustic ray tracing framework.

In summary, the work in this dissertation has produced the first auralizations of the 15th century Vadstena abbey church. In addition, statistical models have been developed, which provides new analytical tools for describing the results of simulations and measurements. Finally, two methods that can be used in room acoustic ray tracing have been developed, one to facilitate simulation of surfaces which scatter sound primarily in one direction, and one to maintain simulation quality when computational resources are limited.

Contents

Part I – Introduction and Contextualization

Part II – Scientific Publications

Paper A
Historically based room acoustic analysis and auralization of a church in the 1470s
Hanna Autio, Mathias Barbagallo, Carolina Ask, Delphine Bard Hagberg, Eva Lindqvist Sandgren, Karin Strinnholm Lagergren
Applied sciences 11.4 (2021), p. 1586. DOI: `10.3390/app11041586`. URL: `https://www.mdpi.com/2076-3417/11/4/1586`

Paper B
The Influence of Different Scattering Algorithms on Room Acoustic Simulation in Rectangular Rooms
Hanna Autio, Nikolaos-Georgios Vardaxis, Delphine Bard Hagberg
Buildings 11.9, (2021), p. 414. DOI: `10.3390/buildings11090414`. URL: `https://www.mdpi.com/2075-5309/11/9/414`

Paper C
A novel algorithm for directional scattering in acoustic ray tracers
Hanna Autio, Erling Nilsson
Acoustics 5.4 (2023), pp. 928-947. DOI: `10.3390/acoustics5040054`. URL: `https://www.mdpi.com/2624-599X/5/4/54`

Paper D
An iterative ray tracing algorithm to increase simulation speed while maintaining overall precision
Hanna Autio, Nikolaos-Georgios Vardaxis, Delphine Bard Hagberg
Acoustics 5.1 (2023), pp. 320-342. DOI: `10.3390/acoustics5010019`. URL: `https://www.mdpi.com/2624-599X/5/1/19`

Paper E

A statistical method for parameter estimation from Schroeder decay curves

Hanna Autio, Delphine Bard Hagberg

In: *Proceedings of INTER-NOISE 2018, Impact of Noise Control Engineering, 26-29 August*. Chicago, IL, 2018.

Part I
Introduction and contextualization

1 Introduction

Sound is ever-present. Sounds continually inform us of our surroundings, from the sound of rain against a windowsill to the echoing space in an empty church or the quiet of our living room. Satisfactory sound environments can contribute to overwhelming experiences in concert halls or cinemas, to intimate moments in for example restaurants, and to restorative times at home. It has been found that noise can have a detrimental effect on overall health [1], but simply the absence of noise does not guarantee that the sound environment is good [2]. By using acoustic simulations, sound environments can be evaluated without the need to perform measurements in a space that may not even exist. In this way, acoustic simulations can be an invaluable tool when designing real spaces.

1.1 Background and motivation

Sounds perceived by a listener are defined by both the sound source and the way it is transmitted, be it by propagation in space or through headphones. When sounds are emitted and propagated through a room, the effects of the transmission can be calculated and estimated using *room acoustic simulation*. Furthermore, the result of that simulation can be made audible through *auralization*, defined as

> *the technique for creating audible sound files from numerical (simulated, measured, synthesized) data*

by Prof. Dr Michael Vorländer [3]. Auralized simulations can be used for example to experience places virtually or to determine whether a planned building is expected to have satisfactory sound environment.

In the world today, more time is spent in virtual sound environments than before. Virtual meetings, telepresence (that is, being digitally present at a remote, real location), virtual reality (VR) experiences and the widespread use of noise-cancelling headphones are all examples of situations where the auditory attention is focused in a space other than the physical space surrounding us. The development has been accelerated due to the global Covid-19 pandemic. In healthcare, VR and telepresence can be used with significant benefits for hospitalized children, elderly people in care homes and others [4, 5, 6]. It has also been shown that virtual visits to new locations can help inform visually impaired individuals of the space, in anticipation of a future real visit [7]. It is also possible to use virtual sound for research into acoustic perception [8]. The sense of presence in VR is an indicator of the effectiveness of the experience [9], when it is used for learning or training, for example. Spatialized sound has been found to increase the sense of presence [10, 11]. When sounds reproduced in VR are intended to emulate some true sound field, it is certainly important to ensure that the reproduced sound field is perceptually close to the truth. Sound simulation is one tool that can ensure that the virtual sound experience is as realistic as possible.

Another application for sound simulation, auralization and VR is related to the concept of cultural heritage. Cultural heritage can be locations, buildings, artifacts or monuments of (among others) historical, symbolic, anthropological or artistic significance[12]. Since 2003, the concept of intangible cultural heritage has been defined by Unesco in the *Convention for the Safeguarding of the Intangible Cultural Heritage* [13]. Intangible cultural heritage includes traditions, practices and expressions that are recognized as part of the cultural heritage of a community or group, and any tangible objects associated therewith. An example of intangible cultural heritage of particular importance to acoustic research is the acoustics of locations that are recognized for their cultural importance and also used for sound-based traditions or rituals. Examples where acousticians have cooperated with researchers from other fields to examine sound environments of particular cultural importance include Notre-Dame de Paris [14], ancient amphitheaters[15] and pre-colonial communication in the Incan empire[16].

The use of auralization and VR within the field of intangible cultural heritage offers new possibilities. The cultural significance of cultural heritage sites means that many people want to visit those locations, either as tourists or to take part in their own heritage practices. However, this is not always possible or desirable. Over-tourism may pose a danger to the preservation of the site [17, 18], the location may be inaccessible, or the location or site may be destroyed or changed beyond recognition. In such cases, a VR visit to the historical location may be an option. When considering such options it is important that the significance of the intangible heritage is respected, by accurately modelling the space including the acoustics.

Over the recent years, there have been several research projects relating specifically to the acoustics of European churches. In Europe, many churches play an important role in the cultural heritage of nations and religious communities due to history or traditions. A comprehensive review article was produced by S. Giron et al. [19], where numerous room acoustic studies in churches were reviewed. In related research, there are a few examples of VR and auralization projects [20, 14, 21, 22], and a few studies of churches in Northern Europe [23, 24, 25, 26]. However, the North European churches studied are quite few, and no auralizations in such places had been performed.

One of the most historically important churches in the Nordic countries is Vadstena abbey church in the south of Sweden. Vadstena abbey was the mother abbey of the Order of the Most Holy Savior, or the Birgittines. The order was founded in 1344 and the associated abbey and church were built later that century. The Birgittines and Vadstena abbey played an important religious, cultural and economic role in the Nordic countries up until the 16th century when the political and religious landscape changed drastically. The church itself was constructed based on the instructions of St Birgitta, but the interior of the building has been changed significantly over the centuries [27, 28]. These changes are expected to have affected the acoustic properties of the church significantly, and the highly reverberant sound field of the modern day is unlikely to represent the historical reality. Vadstena abbey church is a piece of cultural heritage for people in the Nordic countries and for the Birgittine sisters, and a historically based digital reconstruction in VR with auralizations would be valuable.

The inception of this thesis was to work on a historically-based auralization of the sound environment of Vadstena abbey as a whole, in order to produce a comprehensive VR experience. However, over the course of the project it was found that the simulation algorithms were lacking. Consequently, the research effort was channeled into improving the acoustic simulation tools that were suitable for use in this scenario.

As discussed above, acoustic simulations play an important role in creating good VR. Depending on the purpose of the VR experience, there may be different requirements on the simulated sound field, and the quality of the acoustic simulation must be evaluated accordingly. For example, when the VR experience is used for research in acoustic perception of rooms, it is important that the simulated sound field is as realistic as possibly. Conversely, when VR is used as a design tool or as a method for telepresence, it may be more important to ensure that the VR experience elicits the same physical or emotional response in the user. When acoustic simulation is used as a tool for communicating information (such as spatial information for visually impaired users, or the direction and position of enemies in a computer game) spatial information may be more important than other aspects of the sound field. It has been shown that more realistic is not always better [29, 30]. Depending on the purpose, the evaluation tools should vary. It could be expected that some central aspects of the simulated sound field should be accurate, such as the overall decay pattern and the structures of early reflections.

Within a VR context, the virtual environment must respond very quickly to input from the user, and (in the case of portable devices) computational resources may be severely restricted. In addition, VR requires several computationally intense calculations that occur simultaneously. In those cases, visual calculations are often prioritized (with good reason – perceptual lag can lead to motion sickness) and the available resources for audio simulation are critically limited. In those cases, it is necessary that the acoustic simulation method used is fast and has a relatively small memory footprint. Good candidates for such simulation methods can be found in the geometrical acoustics (GA) field.

Of the GA methods, room acoustic ray tracing is arguably the most commonly used. One of the biggest advantages of ray tracing compared to other GA models is that it can include both mirror-like and diffuse reflections, which has been identified as a necessary element in achieving high simulation accuracy [31, 32]. Ray tracing is used in many room acoustic simulation tools [33, 34, 35, 36], but more improvements are needed before the method is fully suited for VR auralization.

In this thesis work, an acoustic ray tracer has been implemented to allow for an in-depth study of the performance of acoustic ray tracers in room acoustic simulation for auralization. The implemented ray tracer has been used to develop and test new algorithms that may improve overall performance.

1.2 Aim and objectives

The work in this thesis has aimed at improving acoustic ray tracing simulations for auralization, especially for VR in cultural heritage. In particular, the goal has been to bridge the gap between a simulated auralization and the true experience of listening to the Birgittine nuns in Vadstena abbey church in the late 15th century.

There are several aspects that should be considered in working towards the specified aim. The ray tracing simulation method itself should be improved, so that important aspects of the sound field are accurately modelled. It is also important that the simulation method is functional under time- and memory restrictions imposed by hardware and intended use cases. Finally, the overall understanding of ray tracing simulations can be improved so that they can be used more efficiently.

In order to improve ray tracing simulations in the specified ways, the objectives below have been identified:

- Implement an acoustic ray tracer on a GPU
- Improve understanding of acoustic ray tracing algorithms and their simulation results by simulating the sound field in Vadstena abbey church in the 15th century
- Based on insights from the previous steps, develop algorithms to improve simulation speed and the models for scattering

The research in this thesis has served to attain the objectives above. In fig. 1.1, the project is illustrated schematically. On the right, the real experience of listening to the Birgittine sisters in the mother abbey is shown, and the research papers related specifically to this topic (Paper **A** and Paper **E**) are marked in the figure. On the left-hand side, an example of the auralization process for VR is shown. A dry sound sample of chant is processed using the results of an acoustic ray tracing simulation and then presented to the users, sounding as if they were in Vadstena abbey church. The objectives presented for this research serves to improve the room acoustic simulation procedure, in order to bridge the gap between the experiences on the right and left-hand sides. Specifically, Paper **B** – Paper **D** present improvements to the ray tracing framework, as marked in the figure.

1.3 Overview

Ray tracer on the GPU

One objective in this work was the implementation of a room acoustic ray tracer, specifically using specialized ray tracing hardware. In fig. 1.1, this work corresponds to the lower left side.

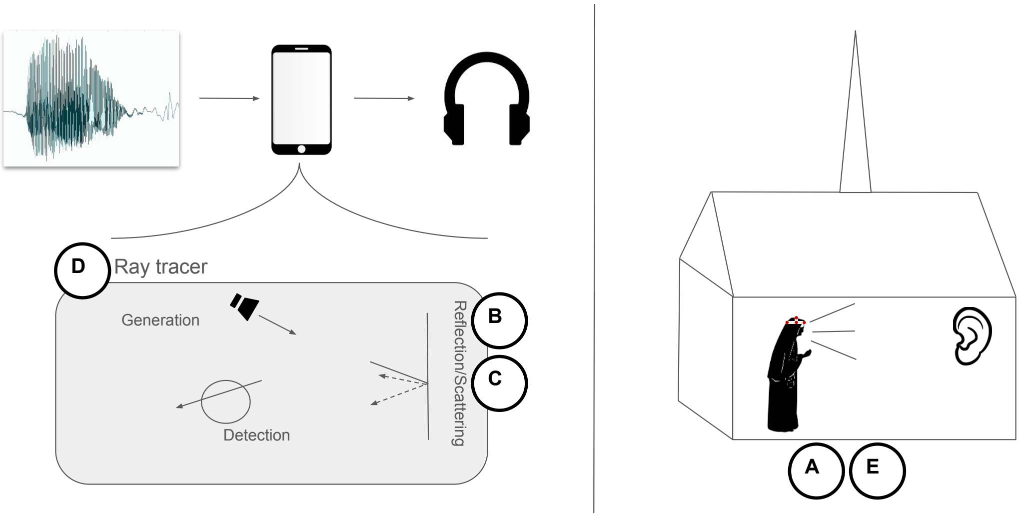


Figure 1.1: Overview of the context of the thesis work. Papers **A**-**E** are marked in the figure. The right shows the auditory experience of listening to a Birgittine nun in Vadstena abbey church, and on the left is the process of auralizing that same experience using a ray tracing simulation. There, a dry sound sample of religious chant is processed using the results of a room acoustic ray tracing simulation and then presented to a user, sounding as if they were in Vadstena abbey church

By using an in-house implementation, it is possible to isolate and test new algorithms explicitly. It is also a good way to ensure that all data from the simulation process can be accessed if needed, either as a tool for analysis or as a tool for debugging.

Using dedicated hardware for acoustic ray tracers is expected to become more and more commonplace, as has been the case for graphics simulation. Consequently, the ray tracer in this work has been implemented to run on such hardware to ensure it is relevant for future developments.

The implementation of a room acoustic ray tracer is also an excellent method of understanding the different parts of the algorithm. This is a necessary element in finding meaningful ways of improving the ray tracer.

The ray tracer implemented in this work is described in detail in chapter 3.

Using simulations of Vadstena abbey to improve understanding of ray tracing simulations

As an initial step in the process of evaluating the current state of ray tracing simulations for auralization, the sound field in a particular location was simulated. The test case was the historically important Vadstena abbey church, indicated on the right of fig. 1.1

A digital historical reconstruction of Vadstena abbey church during the 15th century was performed by a small group of researchers supported by a large community of experts. Absorption coefficients of the floor and ceiling (the only surfaces that are more or less unchanged since the Middle Ages [27]) were calibrated based on reverberation time measurements in the current church, and the digital model was then used for room acoustic simulation using commercial software.

The measurements in the church were subject to large problems, particularly in terms of background noise, as the church was open to the public at the time. The background noise was found to affect particularly impulse response measurements. A statistical model for the measured impulse response and the corresponding Schroeder decay curve under such circumstances was formed and used for reverberation time estimates. The model is described in Paper **E**.

The results of the room acoustic simulation and auralization in Vadstena abbey church are presented in Paper **A**. It is found that the sound field in the church was characterized by the presence of acoustic subspaces with distinct room acoustic properties, one for the monks and one for the nuns of the Birgittine order. The acoustical properties of the subspaces, and particularly the transmission of sound from inside the spaces to the nave, were highly affected by the scattering effects of the vaulted ceiling. Scattering effects are thus identified as a key element in the simulation of Vadstena abbey church.

In the historical Vadstena abbey church, the nave was populated by many objects, such as statues, altars and their decorations, wooden chapels and other objects of religious importance [27, 37]. These objects are expected to scatter the acoustic energy significantly, and scattering in general is thereby identified as an important element again.

Finally, the simulation and auralization were fairly fast to obtain, but yet very far from real-time performance. It was thus clear that methods for improving the speed of simulation were also a necessity for the use of acoustic ray tracers in VR and auralization applications. This is particularly true when it is considered that the calculations at this time were performed on a desktop computer, not a mobile device with much more limited computational capacity.

Two aspects are thus identified of specific relevance to the acoustic simulation of Vadstena abbey church for auralization: Scattering and simulation speed.

Develop algorithms to improve simulation speed and scattering

In this thesis, room acoustic ray tracing simulations have been improved by a solid analysis of current scattering algorithms (Paper **B**), the development of a novel directional scattering algorithm (Paper **C**) and the introduction of an iterative algorithm where the computational power is limited (Paper **D**). In fig. 1.1, each of these papers are marked in connection to the ray tracing simulation on the left.

1.4 Limitations

In this project, the bulk of work has been in the implementation of an acoustic ray tracer, analysis of the simulation results, and on the development of new algorithms. Even so, there are several possible improvements to the ray tracing algorithm that have not been explored. In particular, methods for diffraction [38] should be considered.

Only a limited amount of measurements used for testing have been carried out, and these have been mainly in fairly simple rectangular spaces with little to no furniture. The measurements have been well-suited for testing whether the suggested algorithms can be used, but it should be expected that most real applications are more complex.

The research presented in this thesis is limited to acoustic ray tracing and room acoustics modelling in general, and the goal is to produce simulated sound fields that are *perceptually* similar to the acoustics in real spaces. Accordingly, highly detailed measurements and simulation methods of sound reflection and scattering has not been the focus of this work. Simulation quality is measured by comparing reverberation time, speech clarity and early decay time in the simulated space with measurements.

1.5 Outline of the thesis

In this thesis, the research undertaken to achieve the aims and objectives identified is presented, together with the conclusions it has led to. The main contributions are included in the appended papers, Paper **A** – Paper **E**.

The thesis text is divided into two parts. In Part I, the background of the work is given, together with a brief introduction to the field of room acoustic simulation. In particular, the work of implementing and evaluating a ray tracer is described in some detail. Finally, the main contributions of this work are presented specifically, and some conclusions are discussed. Part II consists of the appended publications, Paper **A** – Paper **E**.

2 Acoustic simulation

Simulation can be defined as

> *something that is made to look, feel, or behave like something else especially so that it can be studied or used to train people.*[39]

Acoustic simulation is thus something that is made to look, feel or behave like sound, often so that it can be studied. It can be used in marine acoustics, medical acoustics such as ultrasound, building acoustics, open air sound propagation and room acoustics. Depending on the field of acoustics and the objectives of the simulation, different elements of the simulation are prioritized and different simulation methods are appropriate.

In many fields, sound and vibrations are used in a very precise way, and it is important that the simulation methods reach a similar level of precision. For example, when using vibrations for non-destructive testing, marine acoustics for mapping the ocean floor or ultrasound acoustics for medical imaging, the instruments used in the field are capable of generating very precise input and output data. The simulation method must reflect this accuracy, and it is important that simulated processes follows the physical reality very closely.

In room acoustics, on the other hand, the degree of accuracy is bounded by the accuracy of the human hearing system. Inside rooms, the spatial resolution needed is typically on the order of magnitude of 1 m [3], and the angular resolution about the head varies from about 1° to over 15° depending on the test signal and the direction [40]. The required resolution for room acoustic simulations is similarly limited.

In this thesis, the main subject has been room acoustics and auralization, specifically using ray tracing. Accordingly, room acoustic simulations and ray tracing particularly will be the focus in this chapter after a brief introduction to core concepts of auralization.

2.1 Auralization

For a proper introduction to the field of auralization, the reader is referred to Vorländer's book on the subject, *Auralization. Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality* [3]. Here, only a brief introduction will be made to familiarize the reader with some key concepts.

As mentioned in the introduction, the term auralization fundamentally refers to the act of making data audible. Within room acoustics, it is used primarily to mean the act of producing a sound sample that sounds as if it were produced and listened to in a specific room. This is done by convolving a dry input signal with the room impulse response for the source and listener positions in the room.

A “dry” input signal refers to a sound recording of only the sound directly generated by the source, no other sounds nor reflections from walls. This can be obtained by using a microphone that is very close to the sound source and sensitive only in a narrow direction, or by recording in an anechoic chamber. In Paper **A**, some recordings in an anechoic chamber are described.

The room impulse response is specific to a given source and listener position within a room, and consists of the sound signal recorded at the listener position when an impulse, or a loud bang, is emitted at the source position. Impulse responses can be measured by firing a start pistol or popping a balloon at the source position and recording the sound at the listener position, which was done in Vadstena abbey within this project (see Paper **E**). It is also possible, and usually preferable, to use methods with more controlled excitation signals such as an exponential sweep, as was done in the measurements in Paper **B**– Paper **D**. Methods for measuring the room impulse response are presented in the international standards ISO 3382-1:2009[41] and ISO 18233:2006[42].

2.2 Room acoustic simulation

Room acoustic simulations can be categorized in one of two classes: Wave-based acoustic modelling, or geometrical acoustics (GA). These differ in the modelling assumptions regarding sound. The two methods are sometimes combined in so-called hybrid programs, such as in Treble [36]. Most commercial acoustic simulation software (such as CATT [43] or Odeon [33]) use GA techniques, reflecting a dominant role in the field since the inception of computer simulation in room acoustics in the 1960s [44, 45]. Even so, there are clear limitations in the GA techniques especially in low frequencies, as evidenced by round-robin studies [46]. Wave-based simulation techniques are more computationally intensive, but are more accurate in terms of sound field modelling.

Wave-based methods derive their name from the wave equation, and aim at finding a solution to it in some form. Sound propagation is governed by the wave equation

$$c^2 \Delta p = \frac{\partial^2 p}{\partial t^2}, \tag{2.1}$$

where p is the local deviation from ambient pressure (i.e. equilibrium) and c is sound propagation speed [47]. Together with boundary conditions and any sources, the wave equation describes a continuous sound field where wave phenomena such as diffraction and resonance follow from the mathematical model. Analytical solutions exist in some special cases, but in most cases approximate numerical methods are used.

There are several methods that can be used to find an approximate solution to the wave equation, and thus an estimate of the sound field in the enclosure. Common methods include the

Finite Element Method (called *FEM*) or Finite Difference Time Domain (*FDTD*) methods. They rely on spatial sampling, and are thus subject to aliasing if the wavelengths in the sound field are too short compared to the sampling distance. These methods are thus more suited to low-frequency sound simulation, as the longer wavelengths allow for a more sparse sampling.

The FEM can be used to solve a wide range of partial differential equations, and some extensions that are useful in other fields can be used also in acoustics. For example, the Discontinuous Galerkin Finite Element Method (first formulated in particle physics) has been adapted for use in room acoustic simulation [48, 49] with good results.

An alternative method is to derive the Helmholtz integral equation from the wave equation, using Green's second identity. This equation is defined over the boundaries of the space, and can be used to find a solution in the frequency domain [48]. Using this formulation requires discretization only of the boundaries, after which the sound field in an arbitrary point can be calculated [1].

Geometrical acoustics, or GA, uses a different and simplified model for sound propagation. It is assumed that the sound wavelengths are small compared to the dimensions of the room, and that the wave fronts follow plane wave behavior [38]. GA methods are significantly less computationally intensive, and work well for higher frequencies. Some GA models are discussed more in detail below.

2.3 Geometrical acoustics

Geometrical acoustics use a simplified model for sound propagation where sound energy travels along straight ray paths, generally disregarding wave-related effects such as diffraction and resonance. GA have been used for sound field prediction long before the introduction of computer-based modelling, and despite some inaccuracies it can be of great practical use [50]. GA were first used for computer modelling in the late 1960s [45, 44], and remain the most popular tool in commercial room acoustic simulation software [47] such as Odeon [33], CATT-Acoustic [43], Pachyderm [51] (for use with Rhinoceros 3D) and EASE [52].

Whereas wave-based simulation techniques are based on the governing equation in the interior volume, GA techniques focuses more heavily on the surfaces of the enclosure and how they interact with the acoustic energy. A detailed overview of different techniques can be found in [38]. In general, it is assumed that the effects of reflection and absorption around the boundaries are necessary and sufficient to determine the acoustics of a room, and that acoustic waves follow simple propagation laws inside a room.

In general, GA models do not include diffraction. Diffraction as a term describes the phenomenon when acoustic energy propagates "around" a corner or edge. The diffracted acoustic energy is not reflected from a distinct surface element, and it is thus outside the realm of the surface-based modelling otherwise used in GA. Even so, there are some extensions to GA techniques that allow for the modelling of diffraction [53], such as the Kirchhoff approximation

[54, 55] and the Geometrical Theory of Diffraction [56] or its uniform correction Uniform geometrical theory of diffraction [57]. A comparison of these methods for use in ultrasonic telemetry can be found in [58]. Diffraction models are implemented in various GA room acoustic simulation tools, for example RAVEN [35, 59], Odeon [33] and CATT-Acoustic [43].

GA techniques are closely related to concepts in computer graphics. Both sound and light can be accurately modelled as wave motion, although the wave properties are typically more evident for sound. This is due to the much larger wavelength of sound waves (from several meters down to a few cm) compared to light waves (about 400 nm - 700 nm). Since GA relies on an assumption that the acoustic wavelengths are small compared to the dimensions of the space, many of the concepts and models used in computer graphics are applicable.

A concept that has been adapted from computer graphics is the rendering equation, in GA referred to as the acoustic rendering equation [38]. It defines how much energy is radiated from a surface element in an enclosure, based on generated and incident energy:

$$l\left(x', \Omega\right) = l_0\left(x', \Omega\right) + \int_G R\left(x, x', \Omega\right) l\left(x, \Gamma\right) \mathrm{d}x. \tag{2.2}$$

In eq. (2.2) above, $l\left(x', \Omega\right)$ is the outgoing time-dependent energy at surface point x' in the direction defined by Ω. On the right-hand side is one term for energy generated by the surface element itself, l_0, and one term for the acoustic energy reflected from the surface. The integral deals with how much energy is reflected as based on how much has been emitted towards it from all other surface elements. As previously defined, $l\left(x, \Gamma\right)$ defines the amount of acoustic energy emitted from x in the direction Γ, which in this context is the direction towards x'. The expression $R\left(x, x', \Omega\right)$ is a reflection kernel, which takes into account all modelled aspects of sound propagation between x and x', such as propagation delay, air attenuation, form factor and visibility etc.

The acoustic rendering equation in GA can be considered an analogue to the wave equation in wave-based modelling, as it describes the flow of acoustic energy. In this sense, it describes the sound field that is the objective of the simulation. The acoustic rendering equation is mathematically less complex, particularly as it is not a partial differential equation and partially because it does not model any effects in the interior of the room, such as diffraction or interference.

The GA method most closely related to the acoustic rendering equation in eq. (2.2) is acoustic radiosity. This model iteratively propagates acoustic energy between surface elements, calculating the acoustic radiation from each surface element in each step based on the irradiated energy in the previous step. Eventually, the radiation of acoustic energy from the surfaces is determined and can be used to form an impulse response for any position in the room. In the most simple version of this model, the surfaces reflect acoustic energy fully diffusely, i.e. independent of the direction of incidence and reflection. The model can be extended to other cases [38]. The acoustic radiosity model is implemented in the simulation program PARISM [60].

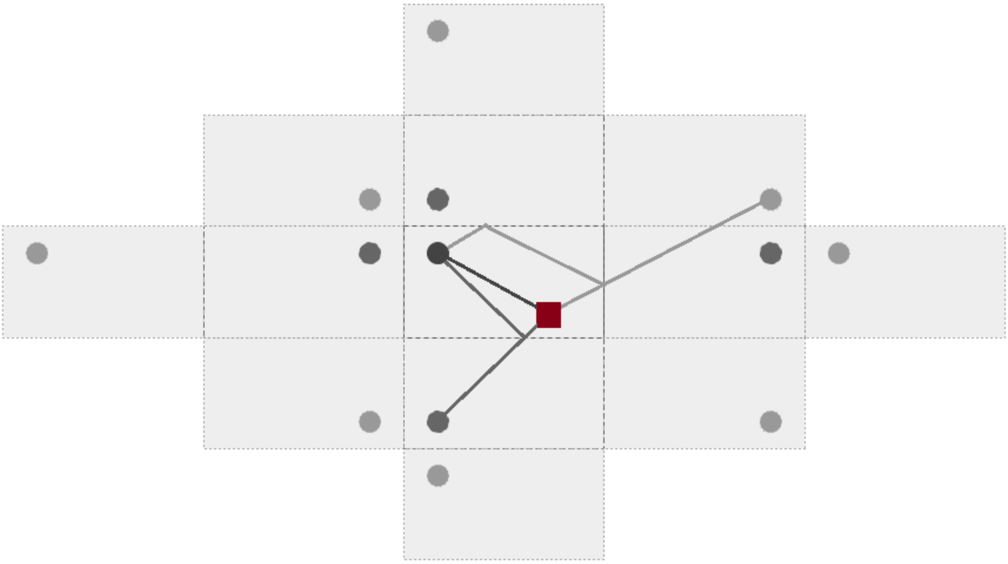


Figure 2.1: Illustration of the virtual sources used in the Image Source Method. The real room is shown in the center, with a square listener and a round source. The image sources that are lighter in color are of higher order, and has undergone more reflections. The direct path, a first- and a second order reflection path are indicated together with the corresponding image sources.

Another common method of GA is the Image source method (ISM). The ISM relies on the concept of virtual sources, similar to what can be experienced visually in mirrors. If sound from a source is reflected in a flat surface in a mirror-like way, the reflected energy can be perceived to originate from a virtual source located behind the wall, where phase and amplitude has been adjusted to account for reflection in the wall. In fact, the energy distributions in the two situations are indistinguishable. The virtual sources are referred to as image sources, and ISM relies on finding all image sources and combining their contributions into an impulse response. A schematic is shown in fig. 2.1, where some image sources (circular) of order up to two is shown. The lighter the source, the higher the order. Note that several image sources will coincide, especially for higher order. The number of image sources increases exponentially with order, and keeping track of them can quickly become very computationally expensive. The ISM is excellent at reconstructing the distinct reflections in the early parts of an impulse response. It is used for this purpose in several room acoustic simulation tools, such as PARISM, EASE or EVERTims [60, 52, 34]. For the later, more reverberant parts of the impulse response there are however two main issues. Firstly, it can be too computationally expensive. Secondly, the assumption that all surfaces behave like acoustic mirrors is too limiting, and some scattering or diffusion must be introduced to achieve a more realistic result [32, 31]

The final GA technique presented here, and arguably the most popular, is acoustic ray tracing.

2.3.1 Acoustic ray tracing

Acoustic ray tracing is a term in GA that defines the specific model used for acoustic propagation, as well as the calculation method used to estimate the predicted sound field. It is used to some extent in many different acoustic simulation tools, such as Treble [36], Odeon [33], RAVEN [59], EVERTims [34], EASE [52], VRWorks Audio [61] etc. In this section, both the underlying physical model and the calculation method are described in some detail.

First, the theoretical model at the core of acoustic ray tracing is considered. In the model, acoustic waves are approximated as a set of discrete sound particles, which travel along straight paths and carry acoustic energy. In principle, an infinite set of particles each carrying an infinitesimal amount of acoustic energy can be used to accurately and fully describe the acoustic energy field inside an enclosure. The model is quite theoretical, but the approximation is useful in explaining and modelling many room acoustic phenomena and can be used both in computer simulations and paper calculations [50]. For example, propagation delay and reflection follow naturally by the particle approximation, and as the particles travel further from a source position they naturally disperse in a way that accurately mimics the reduction of intensity that can be observed when sound energy travels further from a sound source.

Some typical wave phenomena such as diffraction, interference and resonance are not modelled in the ray tracing theoretical framework. The acoustic particles are carriers of acoustic energy and no phase variations are generally modelled, which would be required (to some extent) to account for the phenomena mentioned above. The energetic model is accurate when the wave fronts are mutually incoherent [47]. In general, the ray tracing model is more accurate when the acoustic wavelengths are short compared to the dimensions of the enclosure and the distance between sound source, listener and the walls of the enclosure [47, 3]. When the acoustic wavelength increases, several model elements become less accurate, such as the reflection model, the assumption of mutual incoherence, and the plane wave assumption.

In the ray tracing model, acoustic particles travel along straight paths between the surfaces of the enclosure, like rays. When a particle hits a surface, it is redirected due to reflection and its energy content is adjusted to account for surface absorption. By emitting an infinite set of particles from a source position and using this model for their propagation, an energetic impulse response can be formed by recording each time a particle hits some detector in the listening position. When an infinite number of particles, or rays, are used, the energetic impulse response formed this way is entirely deterministic. When using the model for calculations, the number of particles is by necessity finite. This is discussed more in detail in section 2.3.2.

By discretizing the sound field, acoustic energy can be both reflected specularly and diffusely [32]. This is an advantage compared to acoustic radiosity (where all energy is reflected diffusely) and the ISM (where all energy is reflected specularly). The most common model for ray reflection is that rays are reflected partially specularly and partially diffusely in accordance with Lambert's cosine law [62, 3]. If instead all energy is reflected specularly, the sound field predicted by the ray tracing model and the ISM coincides, and if all energy is reflected diffusely the ray tracing prediction coincides with the radiosity model [38].

The rays or particles are carriers of acoustic energy, rather than acoustic pressure. Accordingly, the ray tracing model produces an *energetic* impulse response, corresponding to the squared impulse response described in section 2.1. An energetic impulse response can be analyzed to obtain a lot of information regarding the perceptual and physical properties of the acoustics, and a statistical model for measured energetic impulse responses is presented in Paper E. However, the energetic impulse response can not readily be used for auralization. There are ways to recover an impulse response from the energetic impulse response [60, 63, 64], but it requires some degree of approximation of the phase response. This topic is not further explored in this thesis.

2.3.2 The ray tracing simulation method

If an infinite set of particles is emitted from a source position and propagated according to the ray tracing model described above, a well-defined and deterministic energetic impulse response is produced. When performing a computer simulation however, an infinite set of particles cannot be considered. Instead, only a randomly selected subset is used. This is in accordance with the Monte Carlo simulation framework, which states that it is possible to form an accurate estimate of a true distribution by sampling only a subset of the population [65]. In practice, this means that a certain number of rays are emitted from the source position, their path and energy content is recorded, and an estimate of the energetic impulse response is formed by recording the particles as they hit a detector in the listener position.

The acoustic ray tracing simulation procedure is thus as follows: A new ray is generated, which has an origin, a direction and some acoustic energy. The ray path, the direction of travel from the origin, is checked to see if the ray hits any surface element. In closed rooms, rays always hit some surface. If the ray hits the ray detector at the listener position, the ray's energy is added to the energetic impulse response estimate at the correct time, and the ray is allowed to travel further along the propagation path. If the ray collides with a wall, the ray energy is adjusted to account for the absorption of the wall and the distance travelled. If the ray energy is below a given threshold after this, the ray is terminated. If not, the ray is reflected and the trace continues, with the position of the wall collision as a new point of origin, and the reflection direction as the new direction of travel.

If a sufficient number of rays, appropriately selected, is traced in this way, the resulting energetic impulse response is an accurate estimate of the energetic impulse response predicted by the underlying ray tracing model. However, the Monte Carlo simulation method is based on randomness and is stochastic in nature. The energetic impulse response produced by the ray tracing algorithm is consequently a stochastic variable which has a statistical distribution, and an expected value and variance. When the algorithm is properly implemented, the expected value of the simulation is the energetic impulse response predicted by the model described above. The variance provides a measure of how much the simulation results are expected to vary between simulation runs.

If it is assumed that the sound energy inside a room is well-mixed, it is possible to form analytical expressions for the statistical properties of the simulated rays. By using these expressions, estimates of the expected value and variance of the simulated energetic impulse response can be produced. Although the assumption of a well-mixed sound field typically does not hold in a given room, the estimates of expected value and variance can be used to understand their interaction and their dependence on various simulation parameters. It is for example possible to use such estimates to determine what number of rays must be used to achieve a certain degree of precision in the simulation results [3], and to estimate the full uncertainty in ray tracing simulation based also on uncertainty in the input parameters [66]. In the work presented in this thesis, new estimates of the expected value and variance have been derived. In these estimates, fewer assumptions are made on the energy levels of each ray. The estimates are presented and tested in Paper **D**.

As is the case with Monte Carlo simulations, it is found that the variance of the estimated energetic impulse response decreases as the number of samples increase. Using a sufficient number of rays is thus necessary to ensure that the acoustic ray tracing simulation result is accurate. Fewer rays are needed to achieve a given precision when the room is small and has uniformly distributed absorption. In Monte Carlo simulations in general, it is also possible to increase precision by clever sampling strategies such as importance sampling [65]. Some comments on this topic are made in chapter 5.

2.4 Scattering in GA

In the discussion of GA so far, the interaction between surface elements and acoustic energy has been identified as a central element and one of the major differences between the different models. One important aspect of that interaction is scattering. Scattering effects have a significant impact on the perceptual sound qualities of a room. When scattering is introduced, it reduces the prevalence of distinct sound reflections and flutter echoes and can counteract resonance. It has been shown that scattering can be used to achieve a better listening experience [67, 62]. In the research presented in this thesis, it was found that the transmission of acoustic energy within Vadstena abbey church was highly affected by the scattering effects of the vaulted ceiling (see Paper **A**). In this context, it is important that the concept of scattering is discussed more thoroughly.

Acoustic scattering has been defined by Morse and Ingaard as the discrepancy between the actual reflection of sound and the idealized reflection of a plane wave in a flat surface [68]. When using the GA framework, all acoustic energy which is not reflected specularly, or absorbed, is scattered. In reality, acoustic waves are reflected from surfaces in complex ways, where the material properties and its geometry interact with the pressure wave front to give rise to diffusion, diffraction and reflection. When the wave front is not plane but spherical, as it is for a true point source, the reflection pattern is yet further from the ideal case considered in GA.

The most commonly used model for scattered acoustic energy is a diffuse distribution. It is assumed that acoustic energy scattered from reflection in a surface is radiated according to

Lambert's cosine distribution,

$$I(\Omega) = I(0)\cos(\Omega). \tag{2.3}$$

In the equation above, $I(\Omega)$ is the intensity emitted in polar angle Ω. With this model, the energy radiated from a surface element is proportional to that element's apparent surface area, which decreases when the viewing angle increases. It should also be noted that the distribution is independent of the direction of incidence. The Lambertian distribution is used frequently in computer graphics for reflected light from rough surfaces. In the context of scattered acoustic energy, it can be considered an appropriate estimate in the absence of more precise data, and its frequent use in GA is motivated.

Reflection of acoustic energy lead to both mirror-like and scattered reflections, and simulation tools need to take this into account. Energy that is not scattered contribute to distinct reflections in the early parts of the impulse response, known as early reflections, which have a central role for speech intelligibility and the acoustic experience [47]. Scattered energy, on the other hand, have been shown to have an important role for the overall perceived room acoustic quality and its inclusion in GA simulations lead to a much improved simulation accuracy [31, 32]. One of the main advantages to acoustic ray tracing is that it can account both for mirror reflections and for diffuse reflections, or a combination thereof.

Scattering can be measured, and there are two coefficients in ISO standards (17497-1:2004 and 17497-2:2012) that are related to measurements of scattering. They are called the scattering coefficient [69] and the diffusion coefficient [70]. Although they both relate to the distribution of acoustic energy after reflection in surface, they are formally and practically different [31]. The scattering coefficient measures the ratio of the reflected energy which is not reflected in a mirror-like way, and thus coincides with how scattering is generally modelled in GA. Note that the scattering coefficient does not take the distribution of scattered energy into account, and it may or may not be similar to a Lambertian distribution in any given case.

The other coefficient related to scattering is the diffusion coefficient [70]. It measures how closely the reflected energy resembles a uniform hemispherical distribution. If the reflected energy precisely matches this distribution, the diffusion coefficient is 1. This coefficient is typically not useful for GA modelling, but can be used to characterize and measure the scattering effects of diffusers [31].

Overall, scattering can be said to increase the disorder in the sound field. High scattering tends to make the sound field more uniform, or more diffuse, simply by redistributing acoustic energy. A possible interpretation of scattering is that it represents the loss of information, when the directional information of the acoustic wave front is lost. With this interpretation, a naive approach to the distribution of acoustic energy after scattering (such as the Lambertian distribution) is sensible – the main goal is to ensure that the new distribution increases disorder in the system. However, this approach is highly inflexible. For a set of phenomena as complex and important as acoustic scattering, more options are needed. In this project, different approaches to modelling scattering in acoustic ray tracing are the main topic of Paper **B** and Paper **C**.

3 Constructing an acoustic ray tracer

In order to test and evaluate ray tracing algorithms, both new and old, an in-house acoustic ray tracer was developed. In this way, almost the whole program could be modified to test new algorithms, adjust implementation choices and tune the overall program. Using an in-house implementation also allowed for more control to isolate specific programming elements, and thus evaluate the influence of different parts of the simulation model and algorithms.

The developed ray tracer uses the *OptiX* ray tracing engine, developed by NVIDIA [71]. The ray tracing engine supplies a framework for the implementation as well as basic functionality for dealing with geometries and ray collision detection. In particular, using OptiX ensures that the ray tracing simulation itself runs on the Graphics Processing Unit (GPU) and that the memory structure for the geometry is properly optimized. Either of these aspects would have required considerable time and effort to implement without the use of OptiX.

As noted, the ray tracer runs on a computer's GPU. The GPU is a specialized piece of hardware that is capable of extreme parallel computing and perform very many small and identical mathematical operations simultaneously. This is useful particularly in fields where matrix operations are frequent, such as artificial intelligence or computer graphics. Modern GPUs often contain hardware optimized specifically for ray tracing, so that it is performed faster and more efficiently. There are several examples of using the GPU to accelerate acoustic simulation [72, 61, 73, 74, 75], and it is expected that the use of ray tracing-ready GPUs for acoustic ray tracing will improve performance[76, 73, 77].

3.1 Description of the program

A schematic illustration of the ray tracer in this work is shown in fig. 3.1, and a more detailed description is given below. In the graph, green rectangles are data structures and purple elements are programs or methods which perform some specific function in the simulation procedure. The ♦ marks elements that have been implemented within this project. Everything contained in the **Tracing engine** block runs on the GPU, and is coded in `CUDA`. Remaining code is in `C++`.

The **Ray tracer class** acts as an umbrella structure and contains several programs and subclasses which are used to set up and perform simulations. All communication with the OptiX ray tracing engine and its associated data structures on the GPU are dealt with by the **Ray tracer class**. It is responsible for reading input data from the user, configuring the ray tracing simulation on the GPU, and returning the energetic impulse response data.

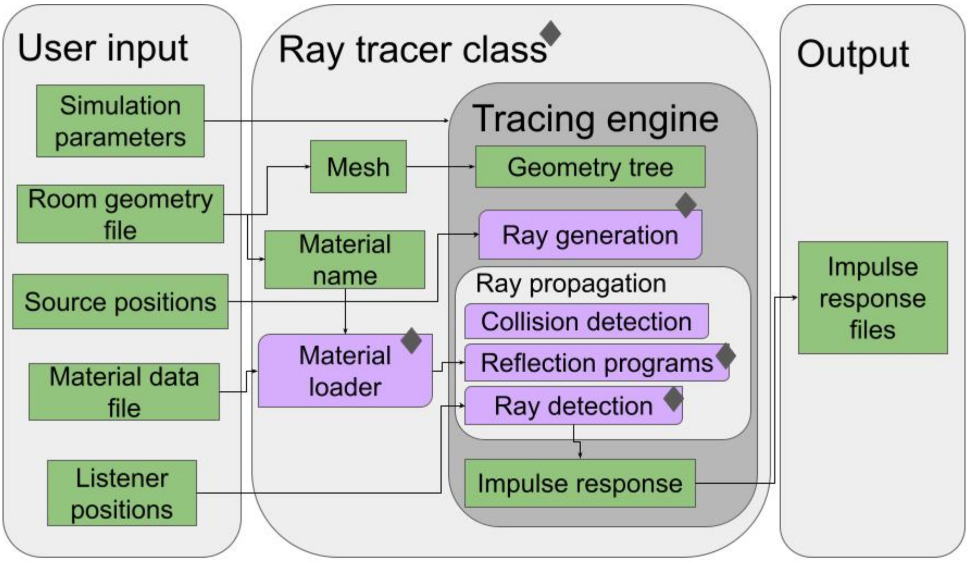


Figure 3.1: Simplified structure of the ray tracer implemented in this project. Purple elements are classes or programs which perform some action, green rectangles are data structures.

3.1.1 Setting up a simulation

When setting up a simulation, the user supplies a JSON[1] data file which specifies the simulation configuration. This file specifies where simulation result data should be written, the number and positions of sources and listeners, the number of rays to be used and other parameters such as sound speed and the unit system used in the geometry file. Also contained in the setup file are the location and name of a geometry file that contains a digital model of the space that should be simulated.

The ray tracer reads the JSON data, and then loads the digital geometry from the specified file using an external library, Assimp [2]. The wireframe model, or mesh, is uploaded to the GPU where OptiX structures it into a geometry tree that allows for fast collision detection. The material data contained in the geometry file is discarded, but the material name is used to refer to a separate material data file.

The material data file is a JSON file which specifies the reflection properties of the material. In the file, absorption and scattering parameters are included, but also the name of the reflection program that should be called when a ray hits that material. Each reflection program implements a specific algorithm for reflection and scattering, and a material loader compiles the reflection programs at run time. In this way, the scattering algorithm can be changed by adjusting the material data file without the need for recompiling the whole program. It is also possible to use different reflection programs and thus different scattering algorithms for

---

[1] JavaScript Object Notation, see https://en.wikipedia.org/wiki/JSON (accessed 2023-11-09)

[2] The Asset Importer Library, see http://assimp.org/ (accessed 2023-11-09)

different materials.

Sound source positions are used to determine the position of ray sources in the ray tracer. The ray sources act as the origin point for the generated rays. This is set as a variable on the GPU, which can be read by OptiX during simulation.

Also contained in the setup JSON file are listener positions and radius. The listeners are modelled as spherical objects and have an associated mesh, contained within the geometry tree on the GPU. The listener geometry is assigned a special program for ray reflections, so that rays that hit the listener are registered in the energetic impulse response and then continues propagation.

When all information has been uploaded to the GPU, the ray tracing simulation can be performed. In this step, rays are generated and propagated in the ray tracing engine based on the ray generation and reflection programs.

3.1.2 Ray generation and propagation

When the acoustic ray tracing simulation is performed, the number of rays prescribed in the setup file is generated and propagated until termination. On the GPU, a large number of rays are simulated simultaneously.

In the ray tracer, acoustic rays consist of an origin, a direction, and a packet of data. Each ray is generated using a ray generation program, which initializes the ray. The origin is the position of the sound source, and the direction is determined by the ray generation program. In the ray generation program used in this project, ray directions are randomly sampled from a uniform spherical distribution. The ray's data packet consists of its energy content (which is initialized to 1), the distance the ray has travelled (initialized to 0) and an index corresponding to the frequency band of the ray. It is necessary to record the frequency of the ray as the reflection programs may treat rays differently depending on their frequency, and a unique energetic impulse response is thus produced for each frequency band.

The same number of rays are used in each frequency (octave or third-octave) band. Since the frequency bands increase in width over the frequency range, this choice has two important implications which are discussed below. Firstly, this leads to a decrease in resolution in terms of rays per frequency, where there are fewer rays per frequency in higher frequency bands. When determining whether this is appropriate, it is important to consider the purpose of the frequency band differentiation. In an acoustic ray tracer, the frequency bands are different because the possible propagation paths differ. Real sound waves interact with surfaces differently depending on their wavelength, and this is shown in part by the distribution of reflected acoustic energy varying across frequencies. In ray tracers, this variation is modelled by rays following different reflection patterns depending on the frequency. Accordingly, the set of possible propagation paths are different for different frequency bands, typically in an octave band resolution. In the ray tracing simulation the rays sample the possible propagation paths, not the frequency range, and the number of rays is thus not related to the range of the frequency bands.

Secondly, the number of rays being equal for each frequency band is related to the amount of energy included in each frequency band. If each ray in the full simulation has the same initial energy, the average energy in each frequency band decreases as the frequency increase. However, this can be accounted for simply by scaling the generated energetic impulse response. As indicated in Paper **D**, scalar multiplication of the energetic impulse response produced from a ray tracing simulation is equivalent to modifying the initial energy of each ray.

The propagation path for each ray is determined by its origin and direction. All surface elements in the propagation path are found, and the one that is closest to the ray origin is identified as the closest hit. The ray reflection program associated with the surface's material is triggered to adjust the ray data in some way and perform any additional tasks. If the ray does not hit any surface, the ray is terminated.

In the reflection program associated with the listener, the ray energy is recorded in the impulse response array. The ray continues its propagation path, but the origin is set to the point of collision with the listener and the distance travelled variable is updated accordingly. The ray direction is not changed. Since the detector is spherical, this procedure makes it so that the ray is inside the detector after collision. In order to ensure that rays are not detected also when exiting the detector, collisions with the listener's inside are ignored. Disregarding the additional detections would not be appropriate if a spatial sampling of the energetic impulse response was performed, but is accurate in the current configuration.

There may be several ray detectors in each simulation. They are each associated with an individual energetic impulse response array on the GPU. In the array, there is a separate energetic impulse response for each frequency band. When a ray is detected, it is recorded in the correct frequency band based on its frequency index. The time of detection is determined by the speed of sound and the distance the ray has travelled, so that the ray energy is recorded at the correct time point in the energetic impulse response array.

In the reflection programs associated with the surfaces of the simulated space, the ray data is updated according to the reflection model used for that material. First, the ray's energy level is adjusted according to the absorption properties of the material and any air attenuation based on the distance travelled. If the ray energy is then below the energy threshold, it is terminated. Otherwise, propagation continues. The ray origin is set to the point of collision, and a new direction is generated based on the reflection algorithm. The distance travelled is updated. The ray can thereafter continue its propagation.

When all rays have been terminated, the simulation is complete. At this time, the process is returned to the ray tracer class which reads the run time energetic impulse responses from the GPU into a local data structure, and then resets the GPU data structure in preparation for the next simulation. The local energetic impulse response data can be returned as a data array for further processing or printed to file.

In order to ensure that the fundamental ray tracing system functions as intended, some simple validation tests were undertaken. A number of basic test cases were set up, where the expected results of a ray tracing simulations could be directly estimated from existing theory. While such tests cannot be used to discern whether the ray tracer can be used for accurate full scale simulations, it is a key step in ensuring that any later simulations are based on a solid model. The overall procedure is described below.

3.1.3 Validation of basic functionality

Two sets of theoretical validation tests were performed, one set to test ray detection and generation, and one set to test ray reflection. Ray detection was tested both in free field and in a digital model of the historical Vadstena abbey church, whereas ray reflection was tested inside a simple rectangular space.

The ray detection tests were designed to evaluate whether rays propagates at the correct speed, to determine if the detected power depends on detector radius and distance to the source as would be expected and to visually evaluate if ray detection works as expected. Some tests were performed in free field, where only a sound source and one or more detectors were modelled. The detectors were positioned in different directions around the sound source, to check whether the random ray generation procedure produces rays in all directions appropriately. No systematic differences between the positions were found.

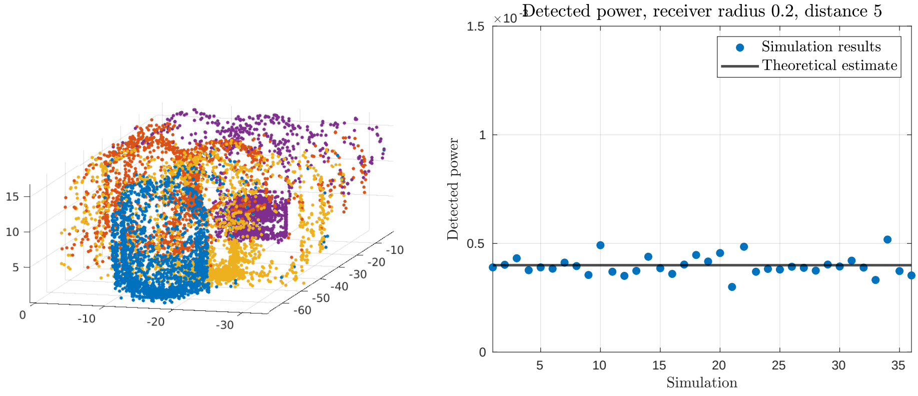
To check propagation speed, an expected time of arrival for the first rays were determined based on sound speed and the distance between source and detector and compared to results from simulations. No discrepancies were found.

Additionally, the geometrical properties of ray detection were tested. A ray tracing simulation was performed for the digital reconstruction of Vadstena abbey church, where the coordinates of the first hit for each ray were recorded. These coordinates are expected to be distributed along the interior surface of the church around the ray source. By plotting the coordinates in a point cloud, an image of the church interior should be formed. The point cloud resulting from one such simulation is shown in fig. 3.2a, where 2000 rays for each of 4 source positions were used. The number of points plotted match the number of rays emitted, indicating that no rays were lost. It is also possible to see the interior structure of the church.

Finally, the total detected power for spherical listeners of different sizes and locations were measured. For an omnidirectional point source, such as the one in this ray tracer, the intensity I at a distance d is determined as

$$I = \frac{P}{4\pi d^2}, \tag{3.1}$$

where P is the emitted power. By integrating this over the surface area of the detector visible from the sound source (corresponding to a hemisphere), a theoretical estimate can be attained for the detected power.



(a) **(b)**

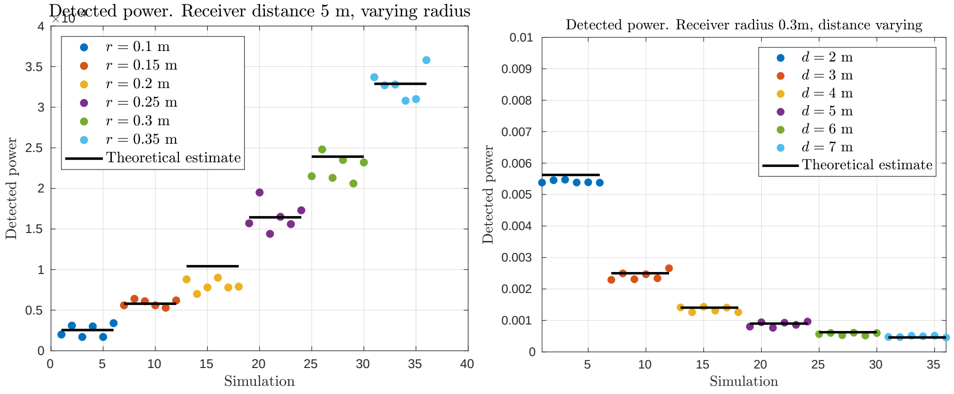
Figure 3.2: Results for some basic validation tests of the ray tracer developed in this project. In fig. 3.2a, a point cloud of the first reflections in a digital recreation of Vadstena abbey church is shown. The different colors indicate different source positions. It is possible to discern the geometry of the interior of the church based on the reflection points, indicating that the simulated rays collide with the surface in an appropriate manner. In fig. 3.2b, the detected energy in a free field simulation is shown, together with a theoretical approximation. In these tests, the detector radius was 0.2 m, and the source and detector distance was 5 m.

The purpose of these tests were primarily to ensure that the detected power varies as expected based on detector size and the distance between the source and the detector, and the precise value of the detected power was of less importance. Accordingly, an approximation of the integration was used, where the total visible surface area of the detector was replaced by its projection, a circular disc of the same radius as the sphere. This disc was irradiated with an intensity corresponding to the intensity at the point of the detector closest to the sound source. This approximation works quite well when the distance between source and detector is large compared to the detector radius. However, it is important to note that it relies on an overestimation of the intensity, an underestimation of the total surface area and completely disregards the effects of the varying viewing angle between the surface elements of the detector and the sound source. Even so, the approximate expression is expected to adequately indicate whether the ray detection works as expected.

A comparison of the detected power as obtained from simulations and as estimated using the method described above is found in fig. 3.2b and fig. 3.3. The results in fig. 3.2b indicate that the detected power overall is quite consistent across simulations, and that it is close to the approximation used. The results in fig. 3.3 additionally show that the total detected energy varies as expected on the distance between source and receiver and on the radius of the detector.

Specifically, the results in this first set of tests are useful in determining whether the random ray generation works properly. Rays seem to disperse in space according to spherical propagation laws, there does not seem to be any error caused by issues in the ray generation, and the propagation speed is as expected.

For the second set of validation simulations, a model of a room was used. The room was rectangular, $(l_x, l_y, l_z) = (3, 4, 5)$m, and all walls had an absorption coefficient of $\alpha = 0.1$. The source was located in $(x_s, y_s, z_s) = (1.5, 3.0, 1.5)$m and the detector in $(x_d, y_d, z_d) =$



(a) **(b)**

Figure 3.3: The results from some free-field simulations in the developed ray tracers. In fig. 3.3a, the receiver radius varies between $r = 0.1$ m and $r = 0.35$ m, with distance $d = 5$ m. In fig. 3.3b, the receiver radius is fixed at $r = 0.3$ m, and the distance varies between $d = 2$ m and $d = 7$ m. The total detected power depends on the distance between source and receiver and on the area of the receiver in the appropriate ways.

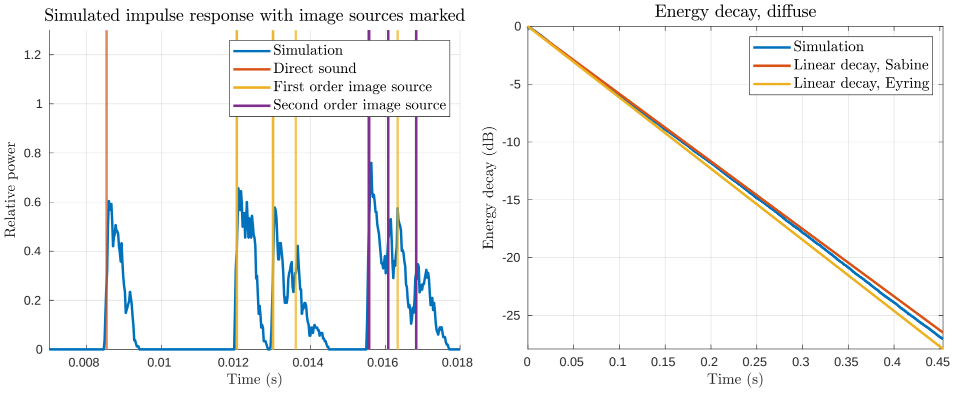
$(1.5, 1.0, 4.0)$m, with a radius of 0.3m. 6×10^6 rays were used, and the sample frequency was 44100 Hz.

The validations from the interior of a room were used to test ray reflection. In one simulation, all rays were reflected specularly. This can be interpreted as a stochastic (and energy based) implementation of the image source model, and the reflection path to arrive at the detector defines an image source. In an image source model (see fig. 2.1), the virtual sources of the lowest order are generally closest to the detector and can thus be seen earliest in the impulse response. They are also relatively few. Accordingly, there should be distinct peaks in the early simulated energetic impulse response corresponding to the image sources. Their time of arrival can be estimated precisely. In fig. 3.4a, the estimated energetic impulse response from one such simulation is shown. The time of arrival of the first image sources are marked in the graph.

As shown in fig. 3.4a, there are peaks in the simulated energetic impulse response for each of the marked image sources. Note that the simulation result has been smoothed (using `MATLABs` built-in function `smooth` with standard setup) to improve visual clarity in the plot. There are some peaks in the simulation result that does not coincide with any of the marked image sources, and these are due to higher order image sources or statistical noise.

In fig. 3.4a, it can clearly be seen that the energy peaks are wider than what would be found in an ordinary image source simulation. This is an effect of the detector being volumetric, and leads to a significant low-pass filtering effect which depends on the detector radius. This should be taken into account when using ray tracing simulations for auralization.

The second validation simulation inside the room tested the effects of diffuse reflection, and whether it could be used to approximate a diffuse sound field. In general, diffuse reflections can cause the sound field to become more diffuse, i.e. have a more evenly distributed energy. In the simulation, all rays were reflected randomly according to Lambert's diffuse distribution.



(a) **(b)**

Figure 3.4: The results of validation tests inside an enclosed space. Left is a simulated energetic impulse response (smoothed from a 44100 Hz sample rate with 5-point moving average), with the time of arrival for the first and second order image sources marked. Right shows the energy decay in a simulation with only diffuse reflections, compared to theoretical estimates of the energy decay in such rooms.

This should lead to a sound field with quite evenly distributed sound energy (or rays) after a very brief period of time. In fig. 3.4b, the energy decay curve is shown in the dB domain. It is linear and smooth immediately at the earliest shown time steps, which is due to the high degree of diffuseness in the sound field.

Although there are no completely diffuse sound fields in practice, there is a rich theoretical foundation to evaluate some of their properties. In particular, the reverberation time can be calculated quite well for a diffuse sound field, using Eyring's or Sabine's formulae [47, 50]. The room used here is expected to have an exponential sound decay curve and a reverberation time of 0.975 s and 1.027 s according to Eyring and Sabine respectively. The simulated reverberation time is determined to 1.002 s, which is within a generally acceptable deviation of 5% from both theoretical estimates. In fig. 3.4b, the energy decay curve in the dB domain is shown, and is linear as expected. Comparison with the decay curves generated from the estimated diffuse field reverberation times from Sabine and Eyring shows that they are very similar.

In conclusion, these tests all indicated that the room acoustic ray tracer implemented the base GA ray tracing model adequately.

3.2 Special considerations related to GPU ray tracing

The GPU is a highly specialized piece of hardware that can perform calculations very efficiently. However, programming for the GPU requires explicit memory handling, and communication between the CPU and the GPU can be relatively time-consuming. If the full potential of the GPU should be utilized, there are many aspects to consider. Some of these, as they relate to acoustic ray tracing specifically, are discussed below.

The advantage of the GPU is primarily its high degree of parallelization. In contrast, memory access is relatively slow and data transfer between the CPU and GPU even more so [76]. As such, the need for memory access and data transfer should be minimized in relation to how many calculations can be performed in each step. Part of this is handled by OptiX itself, in that information is only actually transferred to the GPU when it is needed. An implementation choice that has been affected by this fact is the choice to run the ray tracing step for each frequency band in a single simulation step. Rays of different frequency bands do not interact, and the simulation steps are formally and technically separate. However, by running the simulation for all frequency bands simultaneously the parallelization capabilities are used more fully and all data is transferred at once.

It is also relevant to consider the structure of the reflection programs themselves. On the GPU, different program threads are processed in bundles, or warps, where all threads are subject to the same operations at the same time [76]. This implies that many rays can be grouped and processed on a single subprocessor, where each ray is subject to the same operations at the same time. If there is any branching, for example from an if-statement in the reflection program, the conditional statement will be executed for all threads and thus all rays. The programmer should be careful to ensure that branching structures are avoided if possible. As a practical example, one can take the most common scattering algorithm (presented more in detail in Paper **B**). The algorithm reflects rays either specularly or scatter them. In listing 3.1 below, the algorithm is implemented in two ways: Once with and once without the use of an if-statement. In this case the difference between the two options is small at best, but since this part of the program is triggered every time any ray hits any surface, all improvements can have a significant impact.

Listing 3.1: Two implementations of the most commonly used scattering algorithm

```
//Example one, where an if-statement is used
float random_val = randn(0,1);
if (random_val < material.scattering_coeff) {
    current_ray.direction = random_direction();
} else {
    current_ray.direction = reflection();
}

//Below is an alternative implementation, without the use of the if-
    statement
float random_val = randn(0,1);
// True or False is represented as 1 or 0
float is_scattered = (random_val < material.scattering_coeff);
current_ray.direction = (1 - is_scattered) * reflection() +
    is_scattered * random_direction();
```

There are additional considerations when using ray tracing engines specialized for visual ray tracing for acoustic ray tracing, since there are differences between the propagation properties of light and sound. In particular, light propagates (in practice) infinitely fast, whereas sound has a measurable propagation delay. Consequently, time must be taken into account explicitly and recorded in the simulation result. Since different spaces have different amounts of reverberation and thus different time duration of the impulse response, the memory requirement for the simulation result may differ between simulations. This does not occur in visual ray tracing, and it is possible that the memory allocation for acoustic ray tracing should be more dynamic.

Another difference is that visual ray tracers generally can terminate the simulation faster, i.e. after fewer reflections. This is partially because ray tracing for graphics applications typically serve to simulate highlights and distinct reflections, whereas the more diffuse background light is often generated in advance. This pre-calculation (or baking) can be done by ray tracing as well, but it is not subject to real-time simulation time constraints. In acoustic ray tracing, it is possible to envision a similar methodology where a diffuse sound field is pre-calculated and only some calculations need to be performed in real time, but a larger number of reflections are necessary before the sound field is approximately diffuse. Accordingly, the reflection depth even when considering only distinct reflections or what would be called “highlights” in visual ray tracing is typically greater than what is necessary in visual applications. This is due to the fact that sound waves are not typically as diffusely reflected.

This also means that more rays are needed for acoustic ray tracing. When performing visual ray tracing, it is easier to identify elements that have distinct early reflections and emit light rays towards only those elements. In acoustic ray tracing all surfaces of the enclosure are of interest.

4 Evaluating acoustic ray tracing simulation results

The evaluation of an acoustic simulation tool is a multi-faceted process. In Paper **B**, a methodology for evaluating the performance of an acoustic simulation program is employed. It focuses on three aspects: physical accuracy, simulation speed and usability. In the paper, usability refers particularly to the predictability of the program, and whether small input changes leads to small output changes. If so, the program has high usability. Simulation speed refers to the calculation time, which is especially relevant for real-time applications. Finally, physical accuracy refers to how well the simulated sound field matches a real sound field in the modelled location. This final step is the focus of this chapter.

The term itself, physical accuracy, should be examined and expanded upon slightly. It is described above as how well the simulated sound field matches a real sound field in the modelled location. Importantly, this is evaluated by comparing the simulation *output* to real measurements. In this context, the simulation process itself is considered a black box where some data regarding the geometries of the room is put in and an energetic impulse response is produced. Although the processes and calculations that take place inside the black box are central to the simulation result, they are not evaluated explicitly in this step and the physical accuracy of those algorithms are not considered.

The purpose of room acoustic simulations can vary, and the evaluation method should be informed by the simulation purpose. Often, room acoustic simulations are used as a tool to evaluate the acoustic performance of a planned building. In these cases, parameters can be calculated from the simulation results and compared to values prescribed by the design team or official guidelines. In other cases, simulations are used for research into human hearing and sound perception. In both these cases, it is vital that the simulated sound field can predict the sound field in a corresponding real space as accurately as possible to ensure that the design requirements are fulfilled or that the perceptual research does not become misleading.

Another possible purpose of room acoustic simulations are for use in computer games or VR. In such cases, it may be more important that the simulated sounds evoke a similar physiological or emotional response as the real sound field would. Although it is clear that the sound presented in a VR experience should exhibit accurate spatial information to some extent [3, 11, 10] and that the acoustic effects of the head and ears should be included [78], it has also been shown that increased realism or simulation accuracy does not always increase the sense of realism or presence for the user [29, 9]. Therefor, for VR or similar purposes, there should be a certain degree of physical accuracy in the simulated sound field, but such requirements may only apply to some perceptually important aspects and only to a certain degree.

The typical ray tracing simulation output is an estimate of the energetic room impulse response which is able to fully describe many aspects of the acoustic experience of a listener in the given configuration, but is generally not appropriate as an evaluation tool. It is difficult if not impossible to determine perceptual qualities of the sound field directly from an energetic impulse response, and it is not practical to compare them to each other to deduce similarity. Instead, various single digit parameters are often extracted and used to characterize the sound field. These are referred to as *room acoustic parameters* (RA parameters). In contrast to the energetic impulse response, RA parameters describe only a small aspect of the sound field, but in return they are closely related to how humans perceive said sound field.

4.1 Room acoustic parameters

Room acoustic parameters are typically single-digit values extracted from a room impulse response, which aim at characterizing a perceptually important element of the acoustics in the room. Many different aspects of the sound field can be targeted by the parameters, and common examples are (perceived) reverberance, clarity or spaciousness. RA parameters can be used either to describe the acoustic qualities of a space or as a design tool when planning construction or acoustic treatment. In a construction planning phase, requirements on the acoustical performance may be posed in terms of ranges of RA parameters that are acceptable in the finished space. In those cases, computer simulation is used to predict RA parameters under different circumstances before a design with the required acoustical properties is presented. For this reason, it is important to ensure that an acoustic ray tracer can give accurate estimates of RA parameters. Another reason is to ensure that the perceptual qualities of the simulated sound field are similar to a corresponding real sound field. In this section, the RA parameters used in the project work are presented in more detail.

Room acoustic parameters related to reverberance were first introduced by W.C. Sabine in the late 19th century [79]. They are related to the sound decay process in a space, or how long it takes for the room to become silent after a sound event. The sound decay process can be measured in rooms explicitly, by using noise to excite the sound field until it reaches steady-state, and then abruptly turning the sound source off and measuring the sound level over time. It can also be calculated by backwards integration of the energetic impulse response $h^2(t)$, forming the so-called Schroeder decay curve $EDC(t)$, as shown in eq. (4.1) and fig. 4.1.

The sound decay curve is often approximately exponential or (equivalently) linear in the log (dB) domain. This behavior is consistent with other physical processes where energy flows from a store into a sink[50], such as a warm cup of coffee cooling down to the ambient temperature. Deviations from this behavior are expected particularly when the sound field is non-diffuse, such as early in an impulse response or in spaces where sound absorption is localized to a few highly absorptive patches. When the sound field is diffuse, it can additionally be expected that the energy decay curve is identical in each position in the room.

The most commonly used reverberation parameter and the most commonly used RA parameter overall is the reverberation time T_{60}, as defined by W.C. Sabine and described in ISO 3382-

$$EDC(t) = \int_t^{\infty} h^2(\tau)\mathrm{d}\tau \quad (4.1)$$

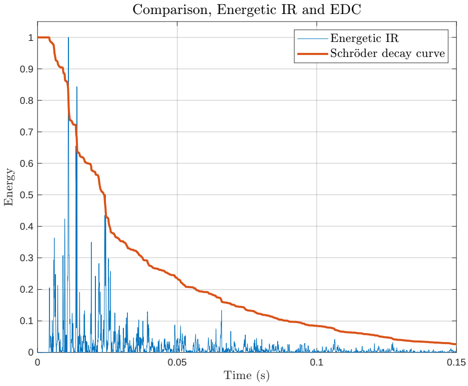


Figure 4.1: Examples of the energetic impulse response and the corresponding Schroeder decay curve, both normalized to have peak value 1. Distinct peaks in the energetic impulse response corresponds to jumps in the decay curve.

1:2009[41]. It is the time needed for the sound level in a room to decrease by 60 dB, or to one millionth of its initial energy. The reverberation time thus clearly takes into account the late parts of the decay process, where the sound field tends to be more diffuse. It is therefor generally expected that the reverberation time is approximately constant over all positions in a room. When evaluating the reverberation time, the first 5 dB of decay is disregarded ([41]), which typically ensures that the level decay curve is more linear by excluding early, non-diffuse reflections. The reverberation time is frequently evaluated by extrapolating from a 20 dB or 30 dB decay process, and can then be referred to as T_{20} or T_{30}. However, it always relates to the time needed for the sound field to decay by 60 dB.

If the decay rate is evaluated from the earliest parts of the decay process, the Early Decay Time (EDT) can be extracted (also described in ISO 3382-1:2009[41]). It is calculated in the same way as the reverberation time T_{60}, but the decay rate from the first 10 dB of decay is used. The EDT is closely related to the early reflections, which tend to be more defined and less diffuse compared to the later parts of the decay process. EDT can thus be expected to vary more between locations. It has been specifically identified as being closely related to the perceived acoustics and reverberation in churches, and thus important for cultural heritage[80].

Another aspect of the perceived acoustics in a space is clarity. Clarity is related to how easy it is to hear and understand speech. It has been found that early reflections are beneficial to the listening experience by increasing the perceived loudness and support of the source sound, whereas late reflections may be perceived as detrimental echoes if they are too strong or defined [47]. Accordingly, clarity parameters tend to be related to the amount of sound energy in the early and late parts of the impulse response.

Some commonly defined clarity parameters are music clarity (C_{80}) or speech clarity (C_{50}), calculated as

$$C_\nu = 10\log_{10}\left(\frac{\int_0^\nu [h^2(t)]\mathrm{d}t}{\int_\nu^\infty [h^2(t)]\mathrm{d}t}\right), \quad (4.2)$$

where $h(t)$ is the impulse response, t is time and ν is a cutoff time between early and late parts of the impulse response, given in ms.

Speech and music clarity thus differs only by the cutoff time used between early and late energy. The different cutoff times reflect a difference in listener preference, as stronger reflections slightly later in the response can decrease speech intelligibility but enhance a music listening experience. It is beneficial to choose the parameter that matches the acoustic activities within a simulated space best, but the two parameters are similar. In this project, speech clarity C_{50} has been favored as this matches most simulated spaces best.

Finally, there are RA parameters that aim at characterizing the spaciousness of the sound field, or the "sensation of space". The concept is difficult to define, but relates to the impression of being enclosed in a large space [47]. It is likely that this should be carefully considered in acoustic simulations for VR, as the goal in these cases is to elicit the sensation of being in a space which corresponds to the virtual environment presented visually. Research has found that the sensation of space is increased when sound energy arrives to a listener from the lateral directions, and when the sound signals on each side are more incoherent [47].

Despite its relevance to VR, no RA parameters related to spaciousness have been considered in this work. This is largely due to limitations in the available measurements, which do not include directional information and thus cannot be used to derive such parameters.

4.2 Comparison to measurements as a tool for validation

Comparison of simulations to measurements is indispensable in terms of evaluating whether the simulation is plausible, but making such comparisons is not trivial. The method for comparing should correctly identify when the data sets are similar or dissimilar in some aspect, but it can be difficult both to determine what should be compared and what constitutes similarity or dissimilarity. When significant dissimilarity is found, identifying the cause of the dissimilarity is often a separate issue that may in turn be very difficult. Comparison between measurements and simulations is necessary, but should be done carefully to ensure that the right things are being compared and that the results of the comparison are meaningful.

In this work, the goal has been to determine whether the simulated energetic impulse response is *perceptually* similar to a real energetic impulse response. It has been determined that RA parameters can be a good tool for this purpose. By studying and comparing RA parameters from simulations and measurements, the physical accuracy of room acoustic simulation results can be evaluated in a way that relates to perceptual similarity.

An important concept regarding RA parameters is the Just Noticeable Difference (JND). The JND is the smallest change to a RA parameter that most humans can perceive. If RA parameters are different but the difference is smaller than the JND, it cannot be expected that a randomly selected listener should be able to hear the deviation. The JNDs generally vary across RA parameters and frequency ranges, and some table values can be found in international standards such as ISO 3382-1:2009[41]. Although it is an important area of research, there are still many questions concerning the JND values for many RA parameters and for different test signals [81, 67].

The accuracy of the measurement data should also be considered. The process of acquiring measurement data should also be done conscientiously, to ensure that the measurements are accurate. Even so, measurements may deviate from reality in some way. Measurement noise or signal processing artifacts are unavoidable to some extent, even when they are minimized by using appropriate measuring methods. Different measurement methods and different parameter extraction methods may yield different results, even when performed immaculately. Room acoustic measurements often rely on signal processing and digital filtering of the recorded data, which may also cause distortion and changes to the data.

The problems associated with measurement accuracy also affects simulation input parameters. In ray tracing simulations, the surface properties of walls and other surfaces are typically included as random-incidence absorption and scattering coefficients. Scattering coefficients can be measured according to the procedures described in ISO 17497-1:2004 [69], but the process itself is complex and time-consuming, and detailed estimates of scattering coefficients are only available for relatively few materials [66]. In addition, the appropriate scattering parameter in a simulation depends on the level of detail in the geometric model of the surface, where a simplified model can be compensated for by increasing the scattering parameter. The random-incidence absorption coefficient can generally be measured more accurately in lab conditions, but the values in situ are difficult to measure and may differ from table values due to mounting variations etc. In all, the input parameters to room acoustic simulations are subject to measurement uncertainty which may affect the result.

All these factors relate to the comparison between simulated and measured RA parameters as a tool for evaluating simulation accuracy. Inaccuracies in the measurement data, both when used for comparison and used as simulation parameters, may introduce variations which do not reflect issues in the modelling software or algorithms. Additionally, the ray tracing model itself is based on approximations of real sound which are only accurate under certain circumstances, and some errors are expected even when both measurements and simulations are performed accurately. Finally, the JND acts as an additional bound on the accuracy that should be expected.

In conclusion, comparison to measurements is an important part of an evaluation scheme. However, evaluation in general is a complex process and this context is no exception. Comparison to measurements is indispensable to evaluate whether the simulation output is realistic, but it is important to be aware of what degree of accuracy is expected and meaningful. Additionally, such comparisons are only one of the tools used for evaluation. Acoustic simulations are used within a context for a specific purpose, and the quality of the simulation program depends also on how easy it is to use in that context. This can be related specifically to simulation speed and memory usage, but also how easy it to set up a fairly accurate simulation.

4.3 Validation of the ray tracer developed in this project

In this project, an acoustic ray tracer has been developed for use in research into ray tracing algorithms for room acoustic simulation, and this is the perspective from which it should be evaluated. Accordingly, it is important that the ray tracer can be used efficiently to try new algorithms and to test the effects of changes in the input data. In addition, it is important that the results obtained from studying this ray tracer can be generalized to other research projects.

The results from the ray tracer used in this project can with high likelihood be extended to other ray tracers. The ray tracer implements the basic ray tracing framework, which is more or less universal in the field. In addition, the ray tracer is implemented for the GPU, which is expected to be in line with future work. These factors ensure that the ray tracer has enough in common with other similar programs that the results apply to those other programs. Importantly, it has also been shown that the ray tracer output is sufficiently accurate to assume that the acoustic ray tracing model is implemented correctly.

The ray tracer developed in this project has been validated by comparing RA parameters derived from simulated energetic impulse responses to measured RA parameters in a few different rooms. In Paper **B**, Paper **C** and Paper **D**, some results from these efforts are shown. In general, the ray tracer developed in this project is reasonably accurate above ca 500 Hz. It predicts reverberation time well, and is usually capable of predicting the EDT and C_{50} to within a few JNDs. The poor results for low frequencies are likely due to limitations in the ray tracing algorithm itself in the modelling of modal behavior. In all, these results indicate that the developed ray tracer implements the ray tracing algorithm accurately and adequately.

5 Extending ray tracing algorithms

All different simulation methods and tools have different advantages and drawbacks, and any number of modifications can be made to adjust the performance of the simulation tool to better suit the purpose of the simulation. Such modifications can range from small changes in the input parameterization to the introduction of completely new modules that can simulate behaviors or effects that were previously outside the scope of the program. In this chapter, larger extensions that somehow modify the behavior of room acoustic ray tracers are discussed.

When extending or modifying a simulation method, the interventions should be aimed at addressing a specific issue or improving a particular facet of the method. This could be for example to reduce simulation noise under certain circumstances, increasing accuracy when modelling a specific element, reducing the memory footprint, or making calculations faster or more precise. It is possible to classify such modifications in many ways, and in this work they are categorized into two groups. The first group is improvements in terms of implementation, either to improve the simulation speed, the memory footprint or reduce errors caused by approximations made in the calculations. The second group consists of improvements in the mathematical model, to account for behaviors that previously have not been modelled accurately or to improve accuracy or precision in the simulation result somehow.

Modifications and extensions frequently cause a worsened performance in some other aspect of the simulation. For example, it is frequently the case that improvements in the mathematical model leads to slower or more memory-intense calculations. Introducing new algorithms to extend the set of phenomena that can accurately be modelled is likely to increase calculation complexity, which is often reflected in an increase in time or memory requirements. Conversely, methods that improve the simulation time or memory footprint often rely on clever approximations which may reduce the output accuracy or precision to some extent. Consequently, it is important to ensure that the modifications lead to an improvement significant enough to outweigh the potential issues they introduce. The benefits and drawbacks should be weighted based on the purpose of the simulation program.

In this thesis, the research has been focused on room acoustic ray tracing for the purpose of auralization, with a particular focus on surface scattering. Suggested improvements should therefor not decrease the suitability of the simulation procedure for auralization, for example by significantly increasing the calculation time. Two extensions to the basic ray tracing algorithm are suggested in this work, described in detail in Paper **C** and Paper **D**. They can broadly be categorized as improvements in the implementation strategy and improvements in the mathematical model respectively. Below, they are put into context with other possible extensions to the ray tracing simulation methodology in each of the categories.

5.1 Improving simulation implementation

The simulation implementation can be improved in different ways. In the case of GPU programming, it is important to ensure that the program exploits the strengths and advantages of the GPU hardware. It has been shown that GPUs can be used to improve simulation performance in acoustic simulations of different types [72, 82]. A good introduction to the topic of acoustics simulation on the GPU is offered in [76]. In this work, the use of CUDA and OptiX as is assumed to ensure that the low-level implementation is appropriate for the GPU in question. Another approach to implementation improvements is presented in [83], where the ray tracing algorithm is sped up significantly by for example source clustering.

An initial improvement to the naive ray tracing approach is to optimize the collision detection between rays and surface. When rays are propagated, all intersections with the surrounding geometry must be found. In the most naive approach, this is done by checking each polygon in the digital model for intersections with the ray path. This can be highly time-consuming for complex geometries with many polygons. In more mature implementations, the geometry is divided into regions such that if a ray path does not intersect a specific region, the polygons in that region does not need to be checked for intersection. Although this may increase memory footprint somewhat, it can significantly reduce calculation time.

Another possible area of improvement, commonly used in other types of Monte Carlo simulations, is to improve the sampling method somehow. This can be used to reduce the number of samples needed to achieve a good simulation result, which can be used to reduce the simulation time. An example is importance sampling, where unlikely events or samples are prioritized to increase likelihood that samples are generated for the whole event space [65]. In acoustic ray tracing, a frequent modification to the sampling strategy is employed in ray generation. Rays are often generated in a pre-determined grid of directions around the source, to ensure that the energy distribution of the sound source is accurately modelled. It is important to note that although the initial ray direction is deterministic in this case, the continued propagation path typically contains random elements due to scattering and the simulation result itself is still stochastic. Even so, this method does reduce the random element and fewer rays should thus be needed to achieve the same precision.

5.1.1 Iterative ray tracing

In this work, a method has been developed which can be used to ensure that the simulation is more robust when the number of samples is reduced. It is presented in more detail in Paper **D**, and there referred to as Iterative ray tracing.

Iterative ray tracing relates to the fact that simulation variance is reduced by the number of samples, or rays, used. If a sufficiently high number of rays is used, the energetic impulse response generated by the simulation method will be identical to what is predicted by the underlying GA model to a very high likelihood. If the number of rays used is lower than that, it becomes more likely that the simulation result should be significantly different. When simulations are used for VR, it is likely that the memory resources and calculation time available will be so restrictive that the number of rays that can be used is limited by what is possible, rather than what is desirable to achieve an excellent result. Iterative ray tracing can be of use in such cases, by increasing the practical number of samples over time through linear combination of several simulation results. This can make it possible to achieve good precision over time in ray tracing simulations in situations where computational resources have previously been insufficient.

5.2 Improving the mathematical model

In this section, some improvements and extensions to the GA model at the core of acoustic ray tracing are discussed.

One example of an area where the GA model can be extended is by including diffraction modelling. There are several approximations of diffraction that can be adapted for use in acoustic ray tracers, and Odeon [33], CATT-Acoustic [43] and RAVEN [59] are all examples of room acoustic simulation software which includes diffraction modelling. There are many different ways of modelling diffraction [58] in GA such as the Uniform geometric Theory of Diffraction (UTD) and the Geometric Theory of Diffraction (GTD) or the Kirchhoff approximation [38]. The effects on the simulation efficiency varies depending on the model implemented, but many of these methods are complex and should be expected to increase calculation time and memory footprint. On the other hand, including a model for diffraction can lead to noticeable improvements in terms of simulation output, for example when line of sight between source and listener is blocked by a screen with diffracting edges. In such cases, the earliest sounds detected by the listener are diffracted around the edges of the screen, whereas the first rays detected in a standard ray tracing simulation will be reflections that may arrive considerably later.

Another example would be to include phase modelling in the acoustic rays, which would improve low-frequency models and simplify auralization of the simulation output significantly. Including phase in the data packet carried by the rays would likely not have a major impact on the simulation performance, but it would introduce meaningful variation between rays which in turn is likely to lead to an increased variance in the simulation output. More rays would be needed to compensate for this. Additionally, absorption would need to be parametrized using complex numbers to account for phase shifts in reflections.

5.2.1 Directional scattering

In this thesis, a method for modelling surface scattering from 1D-acoustic scatterers has been developed, and is presented in Paper **C**. It has been shown that the model can increase simulation accuracy in the presence of 1D-scatterers compared to the uniform scattering model, in particular regarding the late reverberation as exemplified by the reverberation time.

The directional scattering algorithm implements a simple model for sound scattering from 1D-scatterers, and can therefore be implemented without too much efficiency penalties. On the other hand, this also means that the accuracy in terms of the distribution of scattered energy is limited. While the improvements in terms of the reverberation time are promising, the algorithm is likely to be of use mostly in room acoustics as opposed to in more complex modelling systems where explicit diffraction effects are needed.

6 Preview of publications

Paper A

Historically based room acoustic analysis and auralization of a church in the 1470s

Hanna Autio, Mathias Barbagallo, Carolina Ask, Delphine Bard Hagberg, Eva Lindqvist Sandgren, Karin Strinnholm Lagergren

Summary:

Vadstena abbey church is a piece of cultural heritage due to its historical importance as the mother abbey of the religiously, economically and culturally influential Birgittine monastical order. The church in the south of Sweden was built in the 14th and 15th centuries. The interior of the church has been altered significantly since the Middle Ages, both in terms of furnishings and interior construction. In this paper, the processes of developing a digital reconstruction and performing acoustic simulations and auralizations for the historical interior of the church are presented. Religious chants and hymns were recorded for the auralizations, chosen based on historical-musicological research.

Room acoustic analysis found that the sound field in the church was characterized by the presence of two distinct acoustical subspaces in the otherwise highly reverberant nave. One such subspace was located in a choir used by the monks of the religious order. Inside the choir, the religious rites were supported by a much higher music clarity and a less dominant reverberation as compared to the nave. Transmission between the choir and the nave was limited. The second subspace was found on a gallery in the center of the church, designated for the nuns. This location was found to support religious activities by increased clarity and decreased reverberation to an even greater extent than in the choir. In addition, transmission from the gallery into the nave was better. The presence of the acoustic subspaces may reflect a religious intent to separate the congregations from the mundane. Particularly, the acoustic conditions of the nuns were influenced by the central position in the church, which reflected the premiership of the women in the Birgittine order.

A commercial ray tracer was used for the room acoustic simulations. Experiences from this work helped guide the subsequent research in this thesis.

Contributions

The conceptualization of the work and the room acoustic analysis were performed by myself and Mathias Barbagallo. I was responsible for writing the manuscript, with support from Mathias Barbagallo and Carolina Ask. The historical digital recreation of the church was performed by Carolina Ask and Eva Lindqvist-Sandgren. Mathias Barbagallo adapted the historical digital model for acoustic analyses, then performed acoustic simulations and auralizations. The musicological-historical research was performed by Karin Strinnholm-Lagergren. Delphine Bard Hagberg offered supervision. All authors contributed with revisions of the manuscript.

Paper B

The Influence of Different Scattering Algorithms on Room Acoustic Simulation in Rectangular Rooms

Hanna Autio, Nikolaos-Georgios Vardaxis, Delphine Bard Hagberg

Summary:

The differences between three different scattering algorithms used in ray tracers were studied. The three algorithms were: The most common model, with rays randomly reflected specularly or scattered according to a diffuse distribution; A model where all rays are reflected in a direction determined by a combination of random and deterministic elements; A model where, in each reflection, some energy is removed from the ray and added to a "diffuse field" which follows statistical laws for sound decay. In the diffuse field model, all ray reflection is specular. The different scattering methods were discussed in terms of underlying models and the relation to standardized scattering coefficients.

Each of the three algorithms were implemented in a ray tracing program and used for simulation of a rectangular room. The room was measured under two conditions, one highly reverberant and one with an absorptive ceiling. The room was not furnished during measurements or simulations. The three algorithms were used for simulations, and then evaluated based on simulation results, the simulation speed and the usability.

It was found that none of the models could achieve good simulation results in low frequencies in this space, but that the first two models could achieve acceptable results for frequencies from about 1000 Hz and up. It was also found that the most common model was superior in terms of usability, which was evaluated primarily by whether changes in the input scattering parameter leads to predictable variations in the simulation results. Finally, it was noted that the third model was fastest, as it needed fewer rays. In all, it was determined that the most commonly used scattering method works overall best.

Contributions:

The study methodology was developed by me and Nikolaos-Georgios Vardaxis. Measurements were performed by me, as well as algorithm implementation and manuscript writing. Nikolaos-Georgios Vardaxis offered valuable input in analysis and parameter choice. Delphine Bard Hagberg offered supervision, and helped revise the manuscript in cooperation with Nikolaos-Georgios Vardaxis.

Paper C

A novel algorithm for directional scattering in acoustic ray tracers

Hanna Autio, Erling Nilsson

Summary:

The existing algorithms for surface scattering in room acoustic ray tracers typically rely on strong assumptions on the distribution of the scattered acoustic energy, often that it is approximately diffuse according to the Lambertian distribution. In this research project, a new scattering model and corresponding algorithm was presented. It was designed to be of use for 1D-scatterers, such as battened ceilings, stairs or some acoustic diffusers. The scattering algorithm extends the on-off scattering algorithm that was determined in Paper **B** to be overall best. The algorithm is based on a model for an idealized 1D-scatterer, wherein sound energy is scattered into a plane defined by the mirror reflection of the incident ray and the orientation of the scatterer.

The algorithm was implemented and compared to measurements of a space wherein 1D-scatterers had been shown to have a noticeable effect on the room acoustic parameters. It was found that the algorithm could emulate the differences found in the RA parameters, as caused by the orientation of the scatterers. In particular, it outperformed the base on-off scattering algorithm. The parameterization of the algorithm should be studied further.

Contributions

I devised and implemented the algorithm, and developed the scheme for evaluating it. Measurements were obtained from earlier research by a colleague. Erling Nilsson helped with supervision and research guidance, as well as manuscript revisions.

Paper D

An iterative ray tracing algorithm to increase simulation speed while maintaining overall precision

Hanna Autio, Nikolaos-Georgios Vardaxis, Delphine Bard Hagberg

Summary:

Real-time and VR constraints on available memory and calculation time can be restrictive in terms of simulation quality. If, for example, ray tracing simulations must be sped up, the simplest way to achieve this is to reduce the number of rays. However, reducing the number of rays is to reduce the number of Monte Carlo samples, which leads to an increased variance in the simulation output. The simulated impulse response is thus more likely to misrepresent the response predicted by the GA model. In this paper, a method to counteract that effect is introduced. It is called Iterative ray tracing.

In iterative ray tracing, the results from several sets of simulations are combined. In practice, this increases the total number of samples used for simulation, which decreases variance. A statistical model for the results of an acoustic ray tracing simulation is shown, and used to derive an approximation of the expected value and variance of such a simulation. The approximations are validated by comparison to the results of actual simulations, and particularly the relation between simulation variance and the number of rays is studied. It is found that the theoretical estimates follow the same overall trends as the expected value and variance derived from simulation results. The theoretical expressions derived here are useful tools to study the properties of acoustic ray tracing simulations, and in this study in particular they were used to show that it is possible to combine the results from several separate ray tracing simulations to reduce variance. By using specific simulation parameters, the end result can be equivalent to a single simulation with the total number of rays in each constituent simulation. These results are also shown by comparing results from the iterative ray tracing strategy to results from the classical ray tracing method.

Contributions:

The simulation strategy and the statistical model were developed by me. I also performed measurements, implemented the algorithm and performed the analysis. Nikolaos-Georgios Vardaxis assisted with measurements and provided scientific insights during analysis. Delphine Bard Hagberg contributed with manuscript revision and supervision.

Paper E

A statistical method for parameter estimation from Schroeder decay curves

Hanna Autio, Delphine Bard Hagberg

Summary:

Reverberation time measurements and impulse response measurements using impact noise were performed in Vadstena abbey church. At this time, the church was open to the public and in regular use. This lead to significant background noise in the measurements. When reviewing the measurement results, it was found that the measured reverberation time differed significantly depending on whether it was measured directly or derived from the impulse response measurements. The background noise was identified as a possible cause of the discrepancy, as it can have a major impact specifically on the impulse response measurements and the resulting Schröder decay curve.

In this study, a model was formed for the Schröder decay curve consisting of one deterministic and one stochastic part, where the stochastic element describes the influence of the background noise, assuming that the recorded background noise samples are Gaussian distributed and independent. This model is used to derive a likelihood function. A likelihood function gives the likelihood to obtain the given set of measurements as a function of the parameters to a statistical model. The parameters in this case describe the deterministic sound decay. The likelihood function derived in this work can thus be used to determine how likely it is to obtain a certain Schröder decay curve if the true deterministic sound decay follows a specific model. In this project, the sound decay is modelled to be exponential.

The likelihood function is transformed to form the log-likelihood function, which is then optimized to determine which input parameters are most likely to yield the measurement results. When the sound decay is exponential, those parameters can be used to derive the reverberation time.

The method is applied to the measurements from Vadstena. It is found that the reverberation time estimated this way differs from the results from the alternative measurement method and the results from a linear regression model for the energy decay curve.

Contributions:

I performed and analyzed the measurements, developed the model, implemented an optimization scheme and evaluated the results. Delphine Bard Hagberg provided scientific insights, support and manuscript revisions.

7 Concluding remarks

Although ray tracing has been used as a method for room acoustic simulation and calculation for decades, technological advances ensures that continued development and continued research is necessary. As described in the introduction, the research presented in this thesis aims at understanding, evaluating and expanding the room acoustic ray tracing tool set, specifically for use in VR for cultural heritage. To approach that aim, three goals were identified:

- Implement an acoustic ray tracer on a GPU
- Improve understanding of acoustic ray tracing algorithms and their simulation results by simulating the sound field in Vadstena abbey church in the 15th century
- Based on insights from the previous steps, develop algorithms to improve simulation speed and the models for scattering

In the process of achieving these goals, advancements and contributions have been made in several areas. Firstly, the acoustics of the culturally important Vadstena abbey church have been analyzed and recreated. Secondly, the tools for analyzing and describing existing ray tracing methods and models have been extended. Finally, new algorithms have been introduced. Below, the contributions are described in more detail and put into context.

7.1 Contributions

7.1.1 Room acoustics and cultural heritage

The room acoustic properties of the historically important Vadstena abbey church were determined using ray tracing simulations, described in Paper **A**. The simulation results showed that the sound field in the church was characterized by the presence of two distinct acoustical subspaces, one for the monks and one for the nuns. Both of these subspaces exhibited significantly improved conditions for religious services compared to the nave, reflecting the importance of the monks and nuns. In addition, the dominant role of the sisters' congregation was enhanced by better acoustic transmission from their platform into the nave of the church. These results are important from several aspects: They support cultural heritage for the Birgittine order and the public in the region, they may have relevance for other sites with similar religious and monastical traditions, and they are a tool for historical research.

As an element in the reconstruction of the historical Vadstena church, reverberation time measurements were undertaken in the modern church. Although the interior church had changed significantly, the measurements could be used to more accurately estimate the absorption parameters of a significant portion of the equivalent absorption area in the digital model. The data obtained in this measurement endeavor were used to form a new statistical model for the Schroeder decay curve, specifically when background noise significantly affects the measurement data. The model can be used to derive maximum likelihood estimates of energy decay parameters beyond the traditional reverberation time.

7.1.2 Understanding and evaluating acoustic ray tracers

Scattering is often cited as an important feature of room acoustic ray tracers, and in Paper **B** some of the most common scattering algorithms are compared and evaluated. The algorithms are compared based on three different factors; physical accuracy, simulation speed and usability. The structured approach for comparison can easily be extended to additional algorithms and other applications. In the article, the most commonly used scattering algorithm is found to also be overall the best method.

In Paper **B**, the physical models for reflected acoustic energy defined by the scattering algorithms are discussed, as well as their relation to the standardized scattering parameter. In room acoustics in general, it is not necessary to discuss the models in such a detailed way. Instead, the performance of different methods are evaluated on a macro scale, by evaluating the room acoustic parameters produced by the simulation. However, a deeper discussion and a detailed analysis locally at the point of reflection, is expected to be necessary both for improving the methodology for choosing scattering parameters and for developing new algorithms.

As an additional analysis tool, new theoretical estimates of some key statistical properties of room acoustic ray tracing simulations have been derived. In Paper **D**, estimates of the expected value and variance of energetic impulse response generated by room acoustic ray tracing simulation are presented. These estimates are found to be fairly accurate. They can thus be used to for example analyze the effects of changing simulation parameters more closely.

7.1.3 Extending acoustic ray tracing simulations

Two new methods have been developed in this thesis work. The first allows for more complicated scattering models, and the other is used to maintain a high level of precision in ray tracing simulations.

In Paper **C**, an algorithm for directional scattering in room acoustic ray tracers is introduced. The algorithm can emulate changes in room acoustic parameters caused from directional variations in scattering. The modelled energy distribution after reflection is discussed and shown. By the introduction of this algorithm, the acoustic ray tracing tool set is expanded for some situations where previous methods have been unsatisfactory.

Additionally, insights into the statistical behavior of ray tracers were used to develop an algorithm called Iterative ray tracing. Ray tracers are a type of Monte Carlo simulation methods, which fundamentally are more precise when the number of samples increase. Conversely, if too few samples are used due to limitations in the computational resources, the precision is reduced. Iterative ray tracing counteracts such effects by combining the results from several ray tracing steps, thus increasing the practical number of samples (or rays). Iterative ray tracing is presented in Paper **D**, and can be incorporated in existing frameworks with few modifications.

7.2 Future work

The work contained in this thesis can be continued in several ways. Firstly, the approach towards evaluation and improvement of acoustic ray tracing should be used as an inspiration in future work. In order to improve ray tracing simulation beyond the current state-of-the-art, it is necessary to use a focused approach to improve performance in small details but also to evaluate the effects of a modification from a broad perspective. Effects of a change should be considered in terms of simulation results and simulation efficiency, but also in terms of how it affects usage of the simulation tool. The work presented in this thesis, particularly in Paper **B** and Paper **C** exemplifies this approach.

In terms of the algorithms introduced in this work, they should firstly be tested in congruence with existing ray tracing frameworks and for auralization in particular. In doing so, it is likely that ideas for further development arise. Both Iterative ray tracing the directional scattering algorithm introduced here are highly specific in that they affect only a small part of the simulation tool. In this sense, they can be incorporated in a wide range of ray tracers, but also extended in a fairly straight-forward way.

Directional scattering should be evaluated further. Firstly, it should be tested for low frequencies. Secondly, the energy distribution predicted by the algorithm should be compared to measurements. Results from those measurements should be used in congruence with more extensive tests on room acoustic parameters to determine reasonable values for the scattering parameters. It should be noted, however, that the algorithm can be used as-is. Further research is needed to form a strong scientific foundation for the full model, but engineering experience from using the algorithm for real applications is likely to be more important for the future development and use.

7.3 Conclusions

In conclusion, the research project described in this thesis has contributed to the understanding of room acoustic ray tracers and shown how this understanding can be used to improve ray tracing algorithms. Using the approach presented here, it is possible to further improve acoustic simulation in general and acoustic ray tracers in particular.

References

[1] Juan Negreira. “Vibroacoustic performance of wooden buildings”. PhD thesis. Lund, SE: Lund University, 2016.

[2] Nikolaos Vardaxis. “Evaluation of acoustic comfort in apartment buildings”. PhD thesis. Lund, SE: Lund University, 2019.

[3] Michael Vorländer. *Auralization. Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. 2008, pp. 1–335. ISBN: 9783540488293. DOI: `10.1007/978-3-540-48830-9`.

[4] Desert Bus for Hope 2018 and Joe Albietz. *DB2018 - Call-in: Dr. Joe Albietz - Medical Director of Child Life at Children's Hospital Colorado*. [Video] Accessed on 2023-10-25. 2018. URL: `https://youtu.be/LCeLFA4Ngro?si=sCSedUSUKm6y_5MH`.

[5] Daniel Johnston, Hauke Egermann, and Gavin Kearney. “An Interactive Spatial Audio Experience for Children with Autism Spectrum Disorder”. In: *Conference on Immersive and Interactive Audio 2019*. York, GB, 2019.

[6] TT. “Digitala 3D-resor ger guldkant för äldre”. In: *Dagens Medicin* (Dec. 13, 2021). Accessed on 2023-10-25. URL: `https://www.dagensmedicin.se/alla-nyheter/kommunal-halsa/digitala-3d-resor-ger-guldkant-for-aldre/`.

[7] Lorenzo Picinali et al. “Exploration of architectural spaces by blind people using auditory virtual reality for the construction of spatial knowledge”. In: *International Journal of Human Computer Studies* (2014). ISSN: 10715819. DOI: `10.1016/j.ijhcs.2013.12.008`.

[8] Imran Muhammad, Michael Vorländer, and Sabine J. Schlittmeier. “Audio-video virtual reality environments in building acoustics: An exemplary study reproducing performance results and subjective ratings of a laboratory listening experiment”. In: *The Journal of the Acoustical Society of America* 146.3 (Sept. 2019), EL310–EL316. ISSN: 0001-4966. DOI: `10.1121/1.5126598`. URL: `http://asa.scitation.org/doi/10.1121/1.5126598`.

[9] James J. Cummings and Jeremy N. Bailenson. “How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence”. In: *Media Psychology* 19.2 (Apr. 2016), pp. 272–309. ISSN: 15213269. DOI: `10.1080/15213269.2015.1015740`.

[10] Claudia Hendrix and Woodrow Barfield. “Presence in virtual environments as a function of visual and auditory cues”. In: *Proceedings Virtual Reality Annual International Symposium '95*. 1995, pp. 74–82. DOI: `10.1109/VRAIS.1995.512482`.

[11] Sandra Poeschl, Konstantin Wall, and Nicola Doering. “Integration of spatial sound in immersive virtual environments an experimental study on effects of spatial sound on presence”. In: *2013 IEEE Virtual Reality (VR)*. 2013, pp. 129–130. DOI: `10.1109/VR.2013.6549396`.

[12] UNESCO. *Cultural Heritage. Definition*. Accessed on 2023-11-28. 2009. URL: `https://uis.unesco.org/en/glossary-term/cultural-heritage`.

[13] UNESCO. *Convention for the Safeguarding of the Intangible Cultural Heritage*. Accessed on 2023-11-28. 2003. URL: `http://portal.unesco.org/en/ev.php-URL_ID=17716&URL_DO=DO_TOPIC&URL_SECTION=201.html`.

[14] Barteld N. J. Postma and Brian F. G. Katz. “Perceptive and objective evaluation of calibrated room acoustic simulation auralizations”. In: *The Journal of the Acoustical Society of America* 140.6 (2016), pp. 4326–4337. ISSN: 0001-4966. DOI: `10.1121/1.4971422`. URL: `http://asa.scitation.org/doi/10.1121/1.4971422`.

[15] Arianna Astolfi et al. “Measurements of acoustical parameters in the ancient open-air theatre of Tyndaris (Sicily, Italy)”. In: *Applied Sciences (Switzerland)* 10.16 (Aug. 2020). ISSN: 20763417. DOI: `10.3390/app10165680`.

[16] Miriam A. Kolar et al. “The Huánuco Pampa acoustical field survey : an efficient , comparative archaeoacoustical method for studying sonic communication dynamics”. In: *Heritage Science* (2018), pp. 1–25. ISSN: 2050-7445. DOI: `10.1186/s40494-018-0203-4`. URL: `https://doi.org/10.1186/s40494-018-0203-4`.

[17] Evelyn Lau. “Peru shuts parts of Machu Picchu from tourists due to erosion”. In: *The National* (Sept. 29, 2023). Accessed on 2023-10-25. URL: `https://www.thenationalnews.com/travel/2023/09/29/peru-machu-picchu-overtourism/`.

[18] Allan Kaval. “Overtourism, a Venetian curse”. In: *Le Monde* (Sept. 1, 2023). Accessed on 2023-10-25. URL: `https://www.lemonde.fr/en/economy/article/2023/09/01/overtourism-a-venetian-curse_6118491_19.html`.

[19] Sara Girón, Lidia Álvarez-Morales, and Teófilo Zamarreño. “Church acoustics: A state-of-the-art review after several decades of research”. In: *Journal of Sound and Vibration* 411 (Dec. 2017), pp. 378–408. ISSN: 10958568. DOI: `10.1016/j.jsv.2017.09.015`.

[20] Rod Selfridge et al. “Creating Historic Spaces in Virtual Reality using off-the-shelf Audio Plugins”. In: *Conference on Immersive and Interactive Audio 2019*. York, GB, 2019.

[21] Brian F. G. Katz, David Poirier-Quinot, and Barteld N. J. Postma. “Virtual reconstructions of the Théâtre de l ’ Athénée for archeoacoustic study”. In: *Proceedings of the 23rd International Congress on Acoustics*. 1. Aachen, DE, 2019, pp. 303–310.

[22] David Fiala and Vasco Zara. *Dijon 1526 – 3D Saint-Chapel.* Museum exhibit. URL accessed on 2023-10-30. Musée des Beaux-Arts, Dijon, FR, 2014. URL: `https://virtual-music-heritage.fr/portfolio-item/dijon-1526-3d-saint-chapel/`.

[23] Krzysztof Kosała. "Calculation Models for Acoustic Analysis of St. Elizabeth of Hungary Church in Jaworzno Szczakowa". In: *Archives of Acoustics* 41.3 (Sept. 2016), pp. 485–498. DOI: `10.1515/aoa-2016-0047`.

[24] Marek Niemas, Jerzy Sadowski, and Zbigniew Engel. "Acoustic Issues of Sacral Structures". In: *Archives of Acoustics* 23.1 (1998), pp. 87–104.

[25] Nikolay Kanev. "Resonant Vessels in Russian Churches and Their Study in a Concert Hall". In: *Acoustics* 2.2 (June 2020). ISSN: 2624-599X. DOI: `10.3390/acoustics2020023`.

[26] Zbigniew Engel and Krzysztof Kosała. "Acoustic properties of the selected churches in Poland". In: *Mechanics/AGH University of Science and Technology* 24.3 (2005), pp. 173–181.

[27] Iwar Andersson, Sune Ljungstedt, and Gunilla Malm. *Vadstena Klosterkyrka. 1, Kyrkobyggnaden*. Ed. by Erik Cinthio et al. Borås, SE: Riksantikvarieämbetet och Kungl Vitterhets Historie och Antikvitetsakademien, 1991. ISBN: 9171928332.

[28] Hollman, Lennart. *Den heliga Birgittas Reuelaciones Extrauagantes*. Uppsala, SE: Kung. Vitterhets Historie och Antikvitetsakademien, 1956.

[29] Roalt Aalmoes, Merlijn Den Boer, and Henk Veerbeek. "Virtual Reality Aircraft Noise Simulation for Community Engagement". In: *INTER-NOISE 2018, Impact of Noise Control Engineering, 26 – 29 August*. Chicago, IL, 2018.

[30] Richard King, Brett Leonard, and Jack Kelly. "Height channel signal level in immersive audio - how much is enough?" In: *AES 146th Convention, Dublin, Ireland.* Dublin, IE, 2019.

[31] Trevor J. Cox et al. "A tutorial on scattering and diffusion coefficients for room acoustic surfaces". In: *Acta Acustica united with Acustica* 92.1 (2006), pp. 1–15. ISSN: 16101928.

[32] Bengt-Inge Dalenbäck, Mendel Kleiner, and Peter Svensson. "A Macroscopic View of Diffuse Reflection". In: *Audio Engineering Society Convention 95*. New York, NY, Oct. 1993.

[33] ODEON A/S. *ODEON Room Acoustics Software*. [Software]. Lyngby, DK, 2020.

[34] David Poirier-Quinot, Brian F.G. Katz, and Markus Noisternig. "EVERTims : Open source framework for real-time auralization in VR". In: *Proceedings of AM '17, London, United Kingdom, August 23-26*. London: ACM New York, NY, USA, 2017. ISBN: 9781450353731.

[35] Dirk Schröder. "Physically Based Real-Time Auralization of Interactive Virtual Environments". PhD thesis. Aachen, DE: Institute of Technical Acousitcs, 2011.

[36] Treble Technologies. *Treble Technologies – Technical Summary*. Tech. rep. July 2022.

[37] Aron Andersson. *Vadstena klosterkyrka. II, Inredning*. Ed. by Aron Andersson et al. Borås, SE: Riksantikvarieämbetet och Kungl. Vitterhets historie och antikvitets akademien, 1983.

[38] Lauri Savioja and Peter Svensson. "Overview of geometrical room acoustic modeling techniques". In: *The Journal of the Acoustical Society of America* 138.2 (2015), pp. 708–730. ISSN: 0001-4966. DOI: `10.1121/1.4926438`. URL: `http://asa.scitation.org/doi/10.1121/1.4926438`.

[39] "Britannica Dictionary definition of SIMULATION". In: *Britannica Dictionary*. Accessed on 2023-10-24. Encyclopædia Britannica Inc., 2023. URL: `https://www.britannica.com/dictionary/simulation`.

[40] Bosun Xie. *Head-Related Transfer Function and Virtual Auditory Display*. 2nd ed. Plantation, FL: J. Ross Publishing, 2013.

[41] International Organization for Standardization. *Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces*. (ISO standard No. 3382-1:2009). Geneva, CH, 2009.

[42] International Organization for Standardization. *Acoustics – Application of new measurement methods in building and room acoustics*. (ISO standard No. 18233:2006). Geneva, CH, 2006.

[43] Bengt-Inge Dalenbäck. *CATT-Acoustic v9.1 powered by TUCT v2*. [Software]. Mar. 2021.

[44] Asbjørn Krokstad, Svein Strøm, and Svein Sørsdal. "Calculating the acoustical room response by the use of a ray tracing technique". In: *Journal of Sound and Vibration* 8.1 (July 1968), pp. 118–125. ISSN: 0022-460X. DOI: `10.1016/0022-460X(68)90198-3`. URL: `https://www.sciencedirect.com/science/article/pii/0022460X68901983`.

[45] Manfred R. Schroeder. "Digital Simulation of Sound Transmission in Reverberant Spaces". In: *The Journal of the Acoustical Society of America* 47.2A (Feb. 1970), pp. 424–431. ISSN: 0001-4966. DOI: `10.1121/1.1911541`. URL: `http://asa.scitation.org/doi/10.1121/1.1911541`.

[46] Fabian Brinkmann et al. "A round robin on room acoustical simulation and auralization". In: *The Journal of the Acoustical Society of America* 145.4 (2019), pp. 2746–2760.

[47] Heinrich Kuttruff. *Room Acoustics*. 6th ed. Boca Raton, FL: CRC Press, 2017. ISBN: 9781482260434.

[48] Finnur Kári Pind Jörgensson. "Wave-Based Virtual Acoustics". PhD thesis. Copenhagen, DK: Technical University of Denmark, 2020.

[49] Huiqing Wang et al. "Room acoustics modelling in the time-domain with the nodal discontinuous Galerkin method". In: *The Journal of the Acoustical Society of America* 145.4 (Apr. 2019), pp. 2650–2663. ISSN: 0001-4966. DOI: `10.1121/1.5096154`.

[50] Lothar Cremer and Helmut A. Müller. *Principles and Applictions of Room Acoustics*. Barking, GB: Applied science publishers LTD, 1982.

[51] Arthur van der Harten. "Pachyderm Acoustical Simulation: Towards Open-Source Sound Analysis". In: *Architectural Design* 83.2 (Mar. 2013), pp. 138–139. ISSN: 00038504. DOI: `10.1002/ad.1570`.

[52] Ahnert Feistel Media Group. *EASE*. Version 5.58. [Software]. Nov. 26, 2023. URL: `https://www.afmg.eu/en/ease`.

[53] Rendell R. Torres, Peter Svensson, and Mendel Kleiner. "Computation of edge diffraction for more accurate room acoustics auralization". In: *The Journal of the Acoustical Society of America* 109.2 (Feb. 2001), pp. 600–610. ISSN: 0001-4966. DOI: `10.1121/1.1340647`. URL: `http://asa.scitation.org/doi/10.1121/1.1340647`.

[54] Max Born and Emil Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. 7th ed. Cambridge, GB: Cambridge University Press, 1991.

[55] Vasilii M Babich, Mikhail A Lyalinov, and Valery E Grikurov. *Diffraction theory: the Sommerfeld-Malyuzhinets technique*. Oxford, GB: Alpha Science international, 2008.

[56] Vladimir A. Borovikov and Boris E. Kinber. *Geometrical Theory of Diffraction*. London, United Kingdom: The Institution of Electrical Engineers, 1994. ISBN: 9780852968307. DOI: `10.1049/PBEW037E`.

[57] Robert G Kouyoumjian and Prabhakar H Pathak. "A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface". In: *Proceedings of the IEEE* 62.11 (1974), pp. 1448–1461.

[58] Bo Lü et al. "Numerical comparison of acoustic wedge models, with application to ultrasonic telemetry". In: *Ultrasonics* 65 (Feb. 2016), pp. 5–9. ISSN: 0041624X. DOI: `10.1016/j.ultras.2015.10.009`.

[59] Dirk Schröder and Michael Vorländer. "RAVEN: A real-time framework for the auralization of interactive virtual environments". In: *Proceedings of Forum Acusticum 2011 : 27 June - 01 July, Aalborg, Denmark*. Ed. by DAS on behalf of EAA. Aalborg, DK, 2011.

[60] Gerd Marbjerg et al. "Development and validation of a combined phased acoustical radiosity and image source model for predicting sound fields in rooms". In: *The Journal of the Acoustical Society of America* 138.3 (Sept. 2015), pp. 1457–1468. ISSN: 0001-4966. DOI: `10.1121/1.4928297`.

[61] NVIDIA Corporation. *VRWorks - Audio*. [Software]. May 8, 2017. URL: `https://developer.nvidia.com/vrworks/vrworks-audio`.

[62] Trevor J Cox and Peter D'Antonio. *Acoustic Absorbers and Diffusers*. 3rd ed. Boca Raton, FL: Taylor & Francis Group, 2017.

[63] Dirk Schröder, Philipp Dross, and Michael Vorländer. “A Fast Reverberation Estimator for Virtual Environments”. In: *30th AES International Conference on Intelligent Audio Environments*. Saariselkä, FI, 2007, pp. 1–10.

[64] Renate 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”. In: *Applied Acoustics* 38 (1993), pp. 145–159.

[65] Carl Graham and Denis Talay. *Stochastic Simulation and Monte Carlo Methods*. Ed. by P. W. Glynn and Y. Le Jan. Berlin, DE: Springer, 2013. ISBN: 978-3-642-39363-1.

[66] Michael Vorländer. “Computer simulations in room acoustics: Concepts and uncertainties”. In: *The Journal of the Acoustical Society of America* 133.3 (2013), pp. 1203–1213. ISSN: 0001-4966. DOI: `10.1121/1.4788978`. URL: `http://asa.scitation.org/doi/10.1121/1.4788978`.

[67] Emma Arvidsson et al. “Subjective Experience of Speech Depending on the Acoustic Treatment in an Ordinary Room”. In: *International Journal of Environmental Research and Public Health* 18.23 (Nov. 2021), p. 12274. ISSN: 1660-4601. DOI: `10.3390/ijerph182312274`.

[68] Philip M Morse and K Uno Ingard. *Theoretical Acoustics*. New York, NY: McGraw-Hill Book Company, 1968.

[69] International Organization for Standardization. *Acoustics – Sound-scattering properties of surfaces – Part 1: Measurement of the random-incidence scattering coefficient in a reverberation room*. (ISO standard No. 17497-1:2004). Geneva, CH, 2004.

[70] International Organization for Standardization. *Acoustics – Sound-scattering properties of surfaces – Part 2: Measurement of the directional diffusion coefficient in a free field*. (ISO standard No. 17497-2:2012). Geneva, CH, 2012.

[71] Steven G. Parker et al. “OptiX: a general purpose ray tracing engine”. In: *ACM Transactions on Graphics* 29.4 (July 2010). ISSN: 0730-0301. DOI: `10.1145/1778765.1778803`.

[72] Ravish Mehra et al. “WAVE: Interactive Wave-based Sound Propagation for Virtual Environments”. In: *IEEE Transactions on Visualization and Computer Graphics* 21.4 (Apr. 2015), pp. 434–442. ISSN: 10772626. DOI: `10.1109/TVCG.2015.2391858`.

[73] Micah Taylor et al. “Guided Multiview Ray Tracing for Fast Auralization”. In: 18.11 (2012), pp. 1797–1810. DOI: `10.1109/TVCG.2012.27`.

[74] Joacim Stålberg and Mattias Ulmstedt. *GPU Accelerated Ray-tracing for Simulating Sound Propagation in Water*. Linköping, SE, June 2019.

[75] Peter Thoman et al. “Multi-GPU room response simulation with hardware raytracing”. In: *Concurrency and Computation: Practice and Experience* 34.4 (Feb. 2022). ISSN: 1532-0626. DOI: `10.1002/cpe.6663`.

[76] Lauri Savioja, Dinesh Manocha, and Ming C. Ling. "Use of GPUs in room acoustic modeling and auralization". In: *Proceedings of the International Symposium on Room acoustics, ISRA 2010*. Melbourne, AU: International Congress on Acoustics (ICA), Aug. 2010.

[77] Marcin Jedrzejewski and Krzysztof Marasek. "Computation of Room Acoustics using Programmable Video Hardware". In: *Computer Vision and Graphics*. Dordrecht, NL: Kluwer Academic Publishers, pp. 587–592. DOI: `10.1007/1-4020-4179-9_84`.

[78] Ilias Bergstrom et al. "The Plausibility of a String Quartet Performance in Virtual Reality". In: *IEEE Transactions on Visualization and Computer Graphics* 23.4 (2017), pp. 1332–1339. ISSN: 10772626. DOI: `10.1109/TVCG.2017.2657138`.

[79] Wallace Clement Sabine. *Collected papers on acoustics*. Ed. by Theodore Lyman. Cambridge, MA: Harvard University Press, 1922.

[80] Francesco Martellotta. "Subjective study of preferred listening conditions in Italian Catholic churches". In: *Journal of Sound and Vibration* 317.1-2 (Oct. 2008), pp. 378–399. ISSN: 0022460X. DOI: `10.1016/j.jsv.2008.03.014`.

[81] John S. Bradley. "Review of objective room acoustics measures and future needs". In: *Applied Acoustics* 72.10 (Oct. 2011). ISSN: 0003682X. DOI: `10.1016/j.apacoust.2011.04.004`.

[82] Anton Ratnarajah et al. "MESH2IR: Neural Acoustic Impulse Response Generator for Complex 3D Scenes". In: *Proceedings of the 30th ACM International Conference on Multimedia*. MM '22. <conf-loc>, <city>Lisboa</city>, <country>Portugal</country>, </conf-loc>: Association for Computing Machinery, 2022, pp. 924–933. ISBN: 9781450392037. DOI: `10.1145/3503161.3548253`. URL: `https://doi.org/10.1145/3503161.3548253`.

[83] Carl Schissler and Dinesh Manocha. "Interactive Sound Propagation and Rendering for Large Multi-Source Scenes". In: *ACM Trans. Graph.* 36.4 (July 2017). ISSN: 0730-0301. DOI: `10.1145/3072959.2943779`. URL: `https://doi.org/10.1145/3072959.2943779`.

Part II
Scientific publications

Paper A

applied sciences

IMPACT FACTOR 2.7

CITESCORE 4.5

Article

Historically Based Room Acoustic Analysis and Auralization of a Church in the 1470s

Hanna Autio, Mathias Barbagallo, Carolina Ask, Delphine Bard Hagberg, Eva Lindqvist Sandgren and Karin Strinnholm Lagergren

Special Issue
Advances in Architectural Acoustics

Edited by
Dr. Nikolaos M. Papadakis, Prof. Dr. Massimo Garai and Prof. Dr. Georgios E. Stavroulakis

MDPI

https://doi.org/10.3390/app11041586

applied sciences

MDPI

Article

Historically Based Room Acoustic Analysis and Auralization of a Church in the 1470s

Hanna Autio [1,*], Mathias Barbagallo [1,2,*], Carolina Ask [3], Delphine Bard Hagberg [1], Eva Lindqvist Sandgren [4] and Karin Strinnholm Lagergren [5,6]

[1] Division of Engineering Acoustics, Lund University, Box 118, 221 00 Lund, Sweden; Delphine.Bard@construction.lth.se
[2] Brekke & Strand Akustik AB, Box 122, 201 21 Malmö, Sweden
[3] Freelance Consultant, 222 20 Lund, Sweden; caro.ask@gmail.com
[4] Department of Art History, Uppsala University, Box 630, 751 26 Uppsala, Sweden; eva.sandgren@konstvet.uu.se
[5] Department of Music and Art, Linnæus University, 351 95 Växjö , Sweden; karin.strinnholm.lagergren@lnu.se
[6] Alamire Foundation, Leuven University, B-3001 Leuven, Belgium
* Correspondence: hanna.autio@construction.lth.se (H.A.); mathias.barbagallo@construction.lth.se (M.B.)

Abstract: Worship space acoustics have been established as an important part of a nation's cultural heritage and area of acoustic research, but more research is needed regarding the region of northern Europe. This paper describes the historical acoustics of an important abbey church in Sweden in the 1470s. A digital historical reconstruction is developed. Liturgical material specific to this location is recorded and auralized within the digital reconstruction, and a room acoustic analysis is performed. The analysis is guided by liturgical practices in the church and the monastic order connected to it. It is found that the historical sound field in the church is characterized by the existence of two distinct acoustical subspaces within it, each corresponding to a location dedicated to the daily services of the monastical congregations. The subspaces show significantly better acoustic conditions for liturgical activities compared to the nave, which is very reverberant under the conditions of daily services. Acoustic transmission from the two subspaces is limited, indicating that the monastic congregations were visually and acoustically separated from the visitors in the nave and each other.

Keywords: archaeo-acoustics; worship space acoustics; acoustic subspaces; auralization

check for updates

Citation: Autio, H.; Barbagallo, M.; Ask, C.; Bard Hagberg, D.; Lindqvist Sandgren, E.; Strinnholm Lagergren, K. Historically Based Room Acoustic Analysis and Auralization of a Church in the 1470s. *Appl. Sci.* **2021**, *11*, 1586. https://doi.org/10.3390/app11041586

Academic Editor: Nikolaos M. Papadakis

Received: 28 December 2020
Accepted: 5 February 2021
Published: 10 February 2021

1. Introduction

The acoustics of worship spaces are a significant element of a nation's cultural heritage. The concept of cultural heritage is described by UNESCO as

> "those sites, objects and intangible things that have cultural, historical, aesthetic, archaeological, scientific, ethnological or anthropological value to groups and individuals" [1].

Since UNESCO's adoption of the *Convention for the Safeguarding of the Intangible Cultural Heritage* [2] in 2003, the acoustics and acoustical experiences of churches have been established as an important area of research. The interaction between ritualistic and cultural expressions in churches and their acoustics have been the topic of several research projects [3–10]. The comprehensive review article on church acoustics by Girón et al. [11] discusses the efforts of several research teams who have acoustically characterized a large number of churches across the world and, specifically, Europe.

Most of this research has focused on churches located in countries around the Mediterranean sea, resulting in less scientific literature regarding churches in northern Europe. While Polish researchers have made interesting analyses on some churches around the Baltic Sea [12,13] and there is research on some Russian churches [14], there is very little, if any, acoustic research on religious buildings in Scandinavia and, specifically, Sweden. The research presented in this paper sheds new light on intangible cultural heritage in this part of the world by presenting a room acoustic analysis on a digital reconstruction of Vadstena abbey church in Sweden. This work is part of a larger research project [15].

Such *archaeoacoustical* [16] projects, where acoustic simulations in digital reconstructions of historical spaces are performed, have already been undertaken [17–21]. The results of these projects may be combined with visual models to produce Virtual Reality experiences, which has been done by several teams [22–24] and is the goal of this project. Efforts such as these require a tight collaboration between acousticians, 3D artists, historians, and musicologists to tackle the intrinsic multidisciplinary nature of the research. The large amount of heritage objects not yet investigated with such techniques, the relative novelty of the underlying technologies, and the challenges posed by such collaborations motivates further research projects such as the one reported on in this paper.

The combination of archaeoacoustic modeling and visual models often requires some form of *auralization*. Auralization can be defined as "... the process of rendering audible, by physical or mathematical modeling, the soundfield of a source in a space, in such a way as to simulate the binaural listening experience at a given position in the modeled space" [25]. Application and implementation of auralization is in itself a large research area [26–28]. Although auralized audio samples are presented in this paper, the main focus is placed on the development of the 3D model, and on the room acoustic analysis. A foundational course in auralization can be found in [29].

Liturgical practices are characterized by significant auditory elements, such as prayers, chants, or preaching. Understanding of these practices and how they function within a church benefits from room acoustic analyses of churches. Such analyses are based on room acoustic parameters typically computed from impulse responses, which can be estimated using acoustic simulation software. For example, the perception of chant is strongly related to early reverberation [4], and the intelligibility of song and speech can be related to clarity parameters. As within any volume, room acoustic parameters in a church may have a weaker or stronger spatial dependence.

Spatial variations of room acoustic parameters within churches depend partly on varying materials and partly on varying geometry. Churches are often large, complex buildings with vaulted ceilings and spaces such as choirs, transepts, chapels, or apses. This may lead to the formation of acoustic subspaces [30] that are characterized by room acoustic parameters significantly deviating from the rest of the space. For instance, such subspaces may result in varying perceived reverberation as indicated by EDT within the apse, or degradation of clarity far from the chancel.

In some cases, the structural separation of such subspaces serves a liturgical purpose. For example, smaller chapels may be dedicated for funeral procedures, or the choir for the hourly services of a monastic congregation. Distinct liturgical purposes may coincide with distinct liturgical activities, such as the monastical service being characterized by the chanting of the congregation. As the liturgical activities benefit from proper acoustics, the acoustics in a given subspace must be evaluated with respect to the activities at that location. Pedrero et al. [7] presents such research regarding the cathedral in Toledo, which indicates that the acoustical properties in the subspaces supported the activities performed there. Similar results regarding the choir have been seen in several research projects [31–33].

The interaction between religious rites and acoustics is not yet fully understood, especially when spatial variations within churches are considered. It has been established that the acoustic requirements of worship spaces differ from conventional acoustic guidelines [7,34], but there is not yet a strong consensus on what these requirements are. As of yet, the international standard on measurements of room acoustic parameters, ISO 3382, does not contain guidelines for acoustic measurements of worship spaces. Although there exist established methods for such measurements [35] , the lack of an official standard may lead to variations in the measurement methods which in turn cause issues when comparing the results of such measurements. In addition, the lack of guidelines regarding acoustic requirements may cause issues when acousticians are tasked with constructing new worship spaces, or improve existing spaces. The analysis in this paper, where room acoustic parameters within the abbey are combined with a discussion of the religious purposes of various subspaces, may help shed light on these questions.

The goal of this paper is twofold: First, describing the process of constructing a historically accurate archeo-acoustical digital model of an abbey church in Sweden. Second, presenting the analysis of its room acoustic properties. To start, some background information on the abbey is briefly presented. Then, certain aspects of the historically based digital reconstruction are discussed. This description shows that the conclusions of the acoustical analysis rely on sound historical research, and act as inspiration for future, similar projects. Subsequently, recordings of material for auralization are briefly presented. The process of acoustic simulation is described, and the simulated acoustic field is analyzed in detail both globally and locally within acoustic subspaces.

2. Background on the Church

The church targeted for reconstruction in this paper is a Gothic abbey church located in the south of Sweden, built in the 14th and 15th century. It played an important religious and cultural role in the Nordic countries in the middle ages. After the Swedish reformation (1527), the monastery was eventually dissolved and the abbey church fell into disrepair and neglect for almost two centuries. Although some artifacts remain and most of the abbeys extensive library has been preserved, the interior space of the church has been significantly altered due to several renovations [36].

The church itself is oriented west–east, with the chancel in the west. Its nave is divided into three aisles of five bays each, every bay measuring about $11 \times 11\,\mathrm{m}^2$. In the west, the central aisle is extended by an apse of approximately the same size as the bays. The church walls and pillars are built of limestone, and the ceiling vaults are of plastered brick. In the present day, the interior walls are of naked stone, but their rough surface indicate that they were originally plastered. It is also known that plaster was removed during a renovation in the 19th century [36]. A photograph of the modern day church can be seen in Figure 1, and a 3D model of the interior walls and vaulted ceiling is shown in Figure 2.

Figure 1. Present day church. View from east to west.

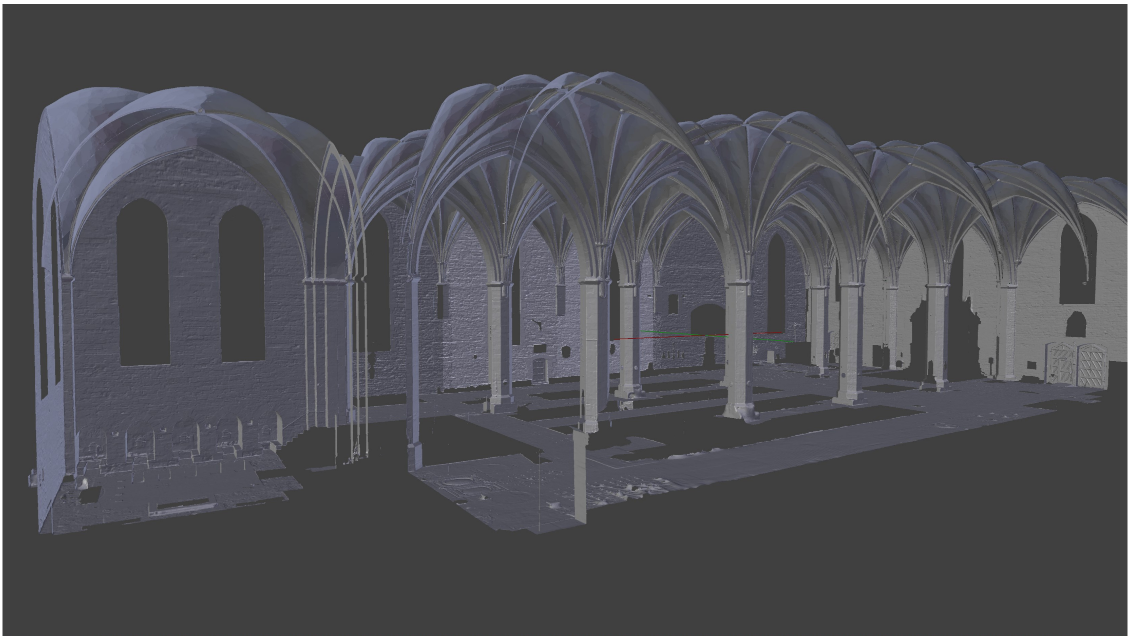


Figure 2. 3D scan showing only the church's original 15th century parts. View to the north, with the chancel and the location of the monks' choir to the west, on the left.

The church and its abbey was the mother abbey of a religious order of both monks and nuns. The monks and nuns lived in separate enclosures, but shared the same church. Both congregations had a specific location within the church where their services were celebrated. The order was characterized by the premiership of women, and is the source of the only known monastical office exclusive for women [37].

The information about the order guides the acoustic analysis to focus on the dedicated spaces for the monastic congregations, especially the nuns'. These two areas are digitally reconstructed, as described in Section 3.1. The monks' choir and the gallery may both be interpreted as choirs. As the choir often has an important role in the acoustics of a church [32,33], a deep acoustic analysis is motivated both by religious and acoustic considerations. However, only traces of the structures remain in the church today, and a historically based digital reconstruction is the only practical option to evaluate the historical sound field of the space.

Below, the processes of digital reconstruction, recording of liturgical elements, and acoustic simulations of the church are presented. The foundations of the historical model are described thoroughly, as one of the goals of the paper is to inspire future researchers in similar projects, and to validate the acoustic results based on historical information.

3. Digital and Historical Reconstruction

The digital reconstruction aims at a time period around year 1470, as this coincides with a period in the abbey's history for which the historical source material is rich. The acoustics of daily religious practices are examined by recording and auralizing material from an ordinary Friday sext. This condition was also the target of room acoustic analysis, characterizing the acoustic properties of the space under normal conditions.

In the following sections, the process of creating the elements for auralization and room acoustic simulation are described. First, the construction of the digital model is discussed. The reconstructions of the spaces for monks and nuns in particular are presented. The recording of material for the auralization is discussed, and finally a few comments are made regarding the adaptation of the high-detail visual model for acoustic simulation software.

3.1. Model Creation

A wide range of historical sources were used in the construction of the digital model. Due to the importance of the church in question, there are many sources directly related to it. This includes drawings, documents, plans, maps, traces in the church room, and historical objects such as sculptures and textiles. In addition, there is a range of earlier research [36,38–42]. This material could be complemented by more general historical information regarding practices for the given time period and geographical location. The collected information was translated into concrete 3D suggestions of what lost and altered parts of the church interior may have looked like.

One of the first steps in the creation of a digital model was the laser-scanning of the complete interior space of the church, aimed at obtaining a high-resolution 3D model to be used as reference. This model was then further processed and refined using the graphical modeling software Blender. First, elements of the scan dating later than the 15th century, as based on historical information, were removed from the digital model. The resulting model of the 15th century shell (Figure 2) was a starting point for the digital reconstruction.

The 15th century shell was extended by 3D suggestions of lost and altered parts of the interior, based on information in historical sources. Different suggestions were then evaluated to find which were more plausible. This process was iterative, and involved primarily the art historian and 3D modeler within the project. Experts on medieval construction, theology, and archaeology were consulted when appropriate.

The evaluation and appraisal of the different 3D suggestions benefited from working in a spatial environment. Formulating the constructions in a 3D space made it clear that some suggestions were incompatible with the church room itself, historical accounts, or other constructions. For example, in some cases several constructions would need to occupy the same space, or they might lack necessary physical support structures. Further refinement could be achieved when liturgical practices were considered. Line of sight to certain spots important for the liturgy, processional walkways, and easy access to key locations was necessary for the religious functionality of the space, and could thus be used to dismiss less appropriate solutions.

The next sections will focus on key areas of the church. First, the furnishings in the nave are discussed briefly. Second, the reconstruction of the gallery and the choir are described and motivated.

3.1.1. The Nave Area

The historical nave of the church was filled with over 60 altars [39] and a multitude of richly decorated grave chapels, giving it a very different impression compared to today (Figure 1). Examples of high-detail reconstructions of chapels and altars are shown in Figure 3. These types of structures are expected to have had a significant impact on the acoustics of the medieval nave. The large number of decorated and complex surfaces lead to increased scattering of the acoustic field, which may have contributed to increased diffusivity. On the other hand, the presence of these structures may have lead to the formation of acoustical subspaces which would imply an increased acoustic heterogeneity.

An additional aspect affecting the acoustic field in the nave was the prevalence of textiles. Textiles played an important part of the furnishing in a late medieval church, and were used as curtains creating small enclosures around most altars (see Figure 3), covers for altars not in use, decorations on altars and walls, and carpets.

The historical reconstruction is expected to reduce reverberation in the nave compared to today, both due to textiles and wood structures such as chapels. The large empty stone volume of the nave today offers very little absorption area, and the reverberation time is long [43]. Wooden grave chapels are expected to increase the amount of absorption across all frequencies, most significantly for the mid-high range. In addition, the large amount of textiles will increase absorption for high frequencies. Together, these effects will lead to a decrease in reverberation across the full frequency range. Such results have been found in other projects [5,44].

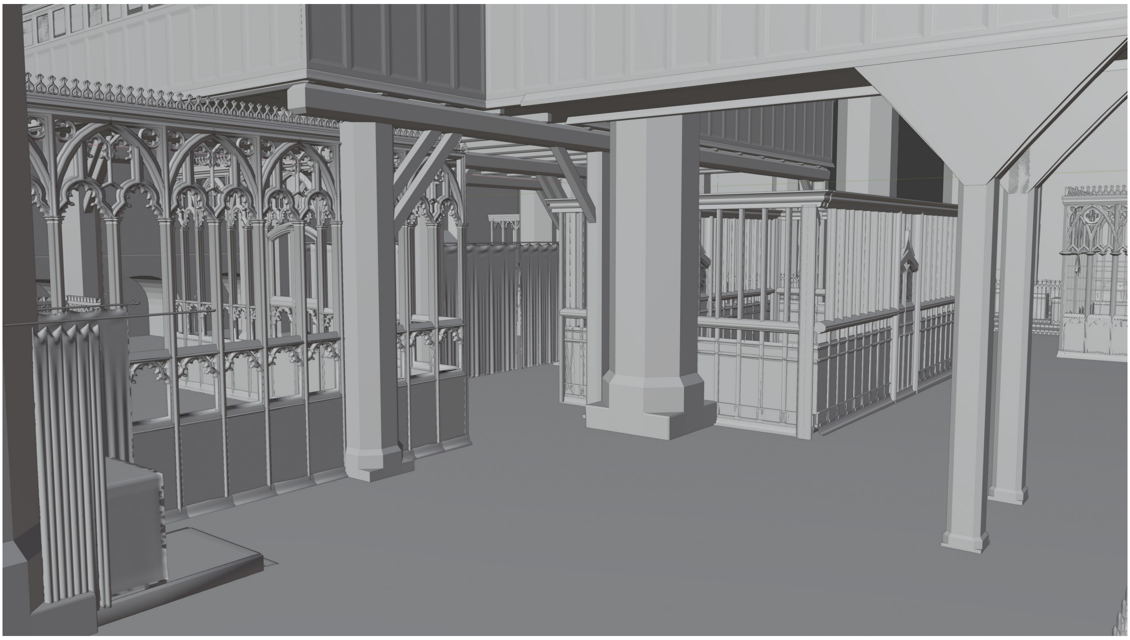


Figure 3. 3D model of three reconstructed chapels (left foreground, middle, right background). In the far left corner is the side of an altar with textile curtains, and a platform on its front covered by a carpet. View from the northeastern side of the nuns' gallery (compare with Figure 4) towards southwest.

3.1.2. The Nuns' Gallery

The construction of an elevated platform for the nuns in the middle of the church is prescribed in the building instructions of the abbey [45]. The only visible remnant of this platform today is a niche near the ceiling vaults that was once its entrance (Figure 4). Research has several suggestions on the specific form and placement of the nuns' gallery [36,38], but there is so far no consensus on its specific location, size, or configuration.

The suggested configuration of the gallery in this project is based on historical documents and physical traces in the church room, and shown in Figure 4. It is surrounded by high panels, and its size was estimated to $10 \times 18\,\mathrm{m}^2$. This is large enough to accommodate the congregation of nuns and the furnishings required for religious purposes [45,46]. In addition, this size allows for the display of a series of paintings along the interior western side of the gallery, which are recorded in [47]. The high panels surrounding the gallery ensure that the nuns could not be seen by anyone else in the church, although see-through lattices allow them visual access to key altars, as required by religious documents [48].

The location of the gallery was determined by using information from a geophysical investigation of the church floor. This investigation revealed anomalies which were interpreted as foundations of support pillars for the gallery. The final model of the gallery, shown in Figure 4, satisfies requirements regarding function and size, is structurally sound and matches the physical traces well.

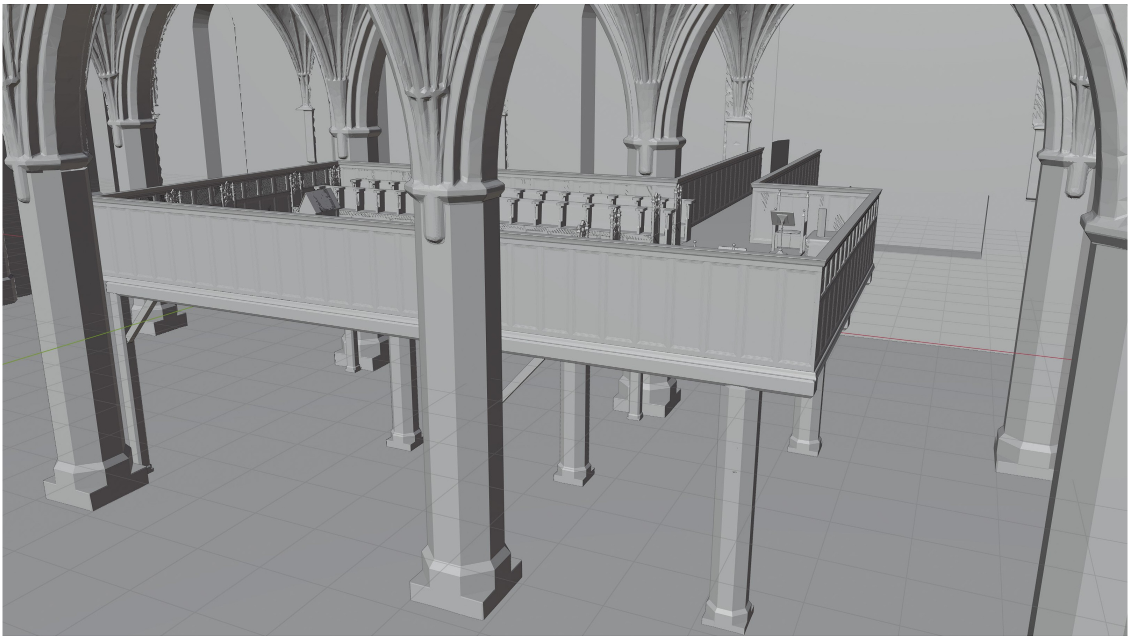


Figure 4. 3D model of the nuns' gallery (other nave elements mentioned in the text are not visualized here), showing the entrance niche in the northern wall, leading through a crossing to the gallery. The panels on the shorter sides have small windows for visual access. The gallery is supported from beneath by four brick pillars. View from south to north.

This interpretation of the gallery is located close to the vaulted ceiling and is separated from the rest of the nave by its floor and panel walls. As such, it likely acts as an acoustical subspace for the congregation of nuns, as is often found in the choir [7,31,32]. As the congregation of nuns gathered there, it is also plausible that the space was characterized by increased absorption by the presence of nuns and their clothing, leading to decreased reverberation and increased clarity. In addition, the gallery's central position close to the vaulted ceiling may have had a significant effect on the transmission of sounds from the gallery.

3.1.3. The Monks' Choir

The monks were located in the choir behind the chancel, in the apse in the west of the church. (left side, Figure 2). As the ground level in this area is about 1 m below the nave, a wooden platform has been suggested to bring the monks to an elevation more liturgically favorable [36]. Possible beam holes and records of the removal of wooden elements from this location support this theory, and it is adopted in this reconstruction.

The shape and size of the wooden platform were determined based on physical traces in the church. Beam holes, book niches, confessionals, and a door opening prescribed the solution shown in Figure 5. Choir stalls were based on duplicates of two choir stalls which remain today.

Similar to the nuns' gallery, this space served the purpose of housing a monastical congregation. Some differences with the nuns' gallery are noted. First, the distance between monks and the vaulted ceiling was significantly larger than that between nuns and the ceiling. Second, the monks' choir was in a recessed position nestled in the apse behind the chancel. It is thus enclosed by close, hard walls on three of four sides, which may affect the sound field within the choir to the point where it differed significantly from that in the gallery and the nave. The acoustic characteristics of this space, especially as compared to the gallery, are further investigated in Section 4.1.

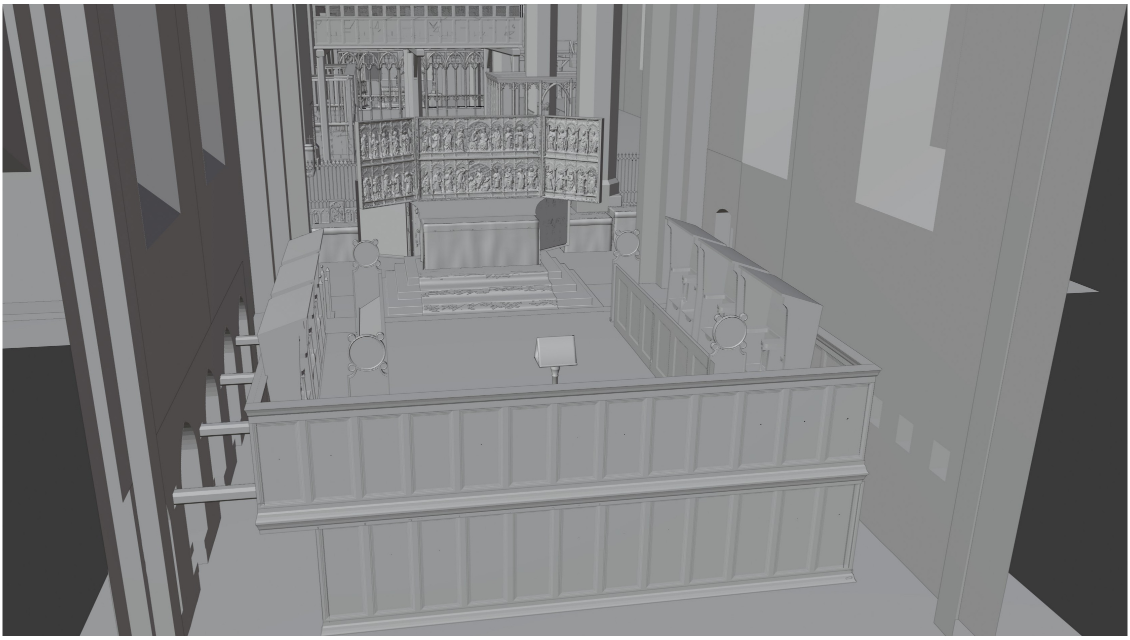


Figure 5. 3D model of the monks' choir with beams fitted in the wall niches to the left, supporting the platform. The back part of the stalls are modeled based on preserved stalls used by the monks. View from the west towards the east. The main altar is visible east of the platform.

3.2. Recording Liturgy

This section describes the overall process of acquiring anechoic recordings of appropriate liturgical material for the auralization task. The full process of selecting material, recording material, and choosing performers will not be fully described here. Only the final choice in material will be presented, and the recordings themselves will be briefly discussed.

The liturgical practices of the order located in this abbey are quite well documented. The daily services consisted of monophonic Gregorian chant, from the respective divine offices of the monks and nuns. The nuns followed their unique liturgical office, as mentioned in Section 2, while the monks observed the liturgy of the diocese where the abbey was located [37]. This information made it possible to recreate a plausible Friday sext, from which elements were chosen for recording. The selected material was chosen to include the important liturgical elements of the service: short responsories, antiphons, prayers, and psalms.

The number of participants in the Friday sext was estimated to about 10 for each congregation. There were to be 60 nuns and 25 monks in the monastery, and only 13 of the monks were ordained and participated in all services. All members of the congregations had responsibilities which may excuse them from services and in particular they might be absent from the small hours (prime, terce, sext, and none). It was concluded that eight male voices and twelve female voices would suffice for the recordings.

3.2.1. Recordings

The recordings were made in the anechoic chamber at Engineering acoustics, LTH, Lund University. Four male and three female singers, familiar with the style of music and the material, were recruited for the recordings. They were recorded using four close mics model Milab VM-44, hanging from the walls.

In order to reach a plausible number of voices in the recordings, the singers were recorded multiple times. During each recording, all participants belonging to the same group were performing together in the anechoic chamber. One of them was wearing headphones, playing back any earlier recordings and fed with generic live 6 s reverb effect; this time length was chosen as the singers deemed it more helpful to achieve good results as compared to no, or 3 s, reverb. The others followed this individual. The variations caused by this method are thought to be consistent with the assumption that these individuals were performing a daily task, albeit as part of a service. Recording and mixing were done in Cubase 11.0.

Singing in a anechoic chamber may pose challenges for singers both due to lack of support and response from the room, and due to the physical influence such environment may have on humans. Such circumstances have an impact on the performance itself, and may result in a presentation that is quite different from how the same material would be presented in a more traditional space, such as a church. The addition of artificial reverb, the choice of singers, and great care during the recording process were used to combat these effects, but it is not possible to guarantee that the presentation is not affected by the discrepancy between the recording conditions and the church. Such problems have been encountered in other research projects, and have not yet been solved [22,49].

A musicologist familiar with the material was present for the full duration of the recordings, ensuring that the performance was as accurate as possible. As the act of performing in an anechoic chamber can be very challenging, the singers took frequent breaks. During these breaks, the performers, the producer, and the musicologist listened together to the recording, both dry and with added reverb, and determined whether it was of an appropriate tempo and quality.

In addition to elements of the sext, some additional sounds were recorded. These were background sounds including prayers, said by female and male voices, the sound of historically representative clothing, sounds from a rosary, sounds from walking with leather shoes against a stone surface, and coughing.

3.3. Room Acoustic Simulations

The digital model constructed in Section 3.1 was exported from Blender to Google Sketch-up and therein adapted for simulations in ODEON 16.0 [50]. Simulations were performed both to characterize the historical sound field in the reconstructed abbey and with the goal of producing auralizations of the sext itself. This section describes the process of adapting the digital model for simulation, as well as the choices made for simulation and auralization.

During the simulation, air absorption was tuned to conditions of 18°C and 50% relative humidity.

3.3.1. Adaptation of the Digital Model

The high-detail model used for visual presentation was transformed into a digital model for acoustic simulation, primarily by simplification of various surfaces. The structure of the ceiling vaults was simplified significantly, and detailed models of sculptures and altar decorations were replaced by simple geometric forms. This improves the performance of the acoustic simulations with regards to accuracy and to calculation speed. Examples of simplifications are shown in Figures 6 and 7. The simplified model is exported from Google Sketch-up using ODEON's exporting tool and imported to ODEON. The exported model counts 5600 surfaces and takes around 1.5 s to export to ODEON format.

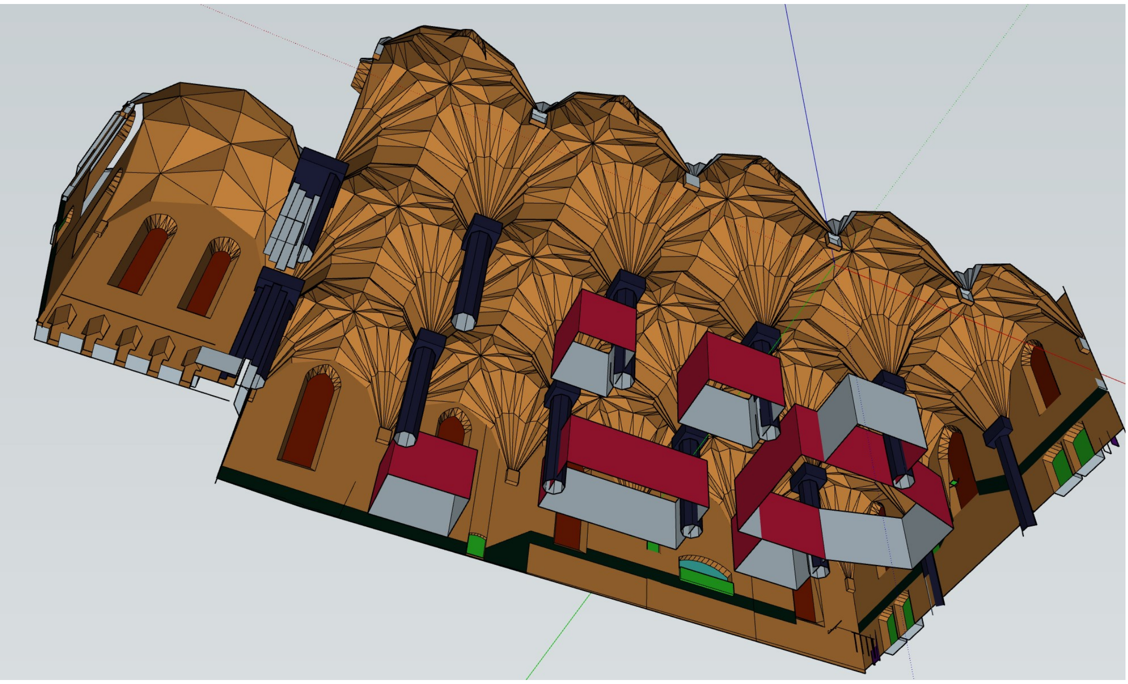


Figure 6. Simplification of vaults and chapels in the acoustic model, view from below with floor layer being hidden. The elaborate wooden walls around the altars and graves (compare to Figure 3) are rendered with boxes which are assigned a 45% transparency in ODEON. Screenshot from Google Sketch-up.

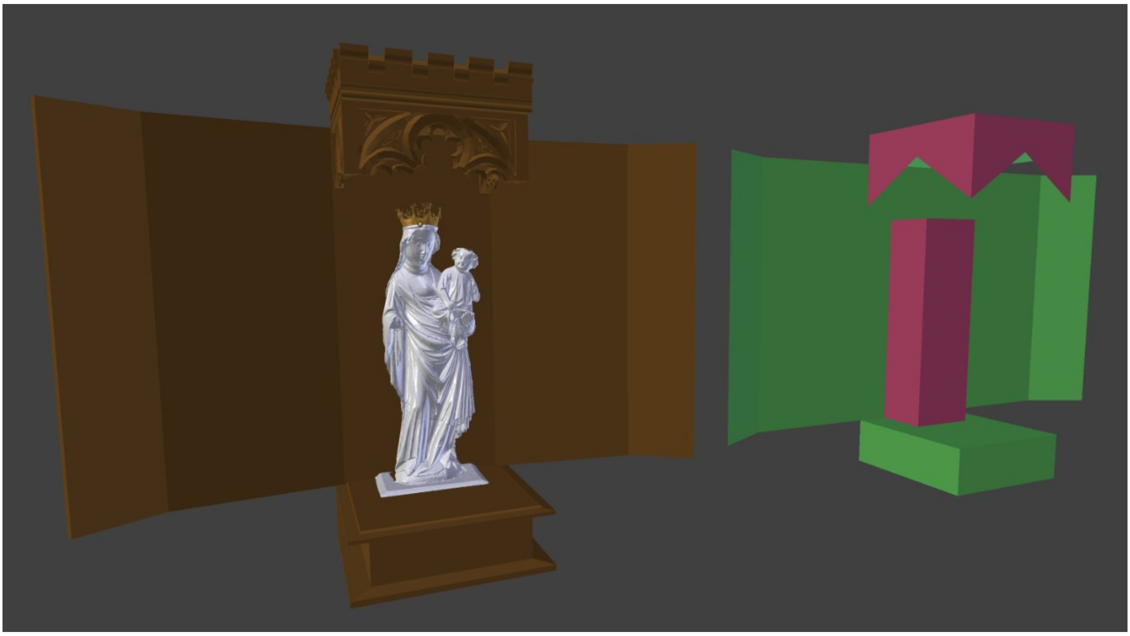


Figure 7. Simplification of statues from the visual model to the acoustic model. Screenshot from Google Sketch-up.

3.3.2. Acoustic Properties of Materials in the Church and Calibration of the Model

The acoustic properties of the various materials in the church needed to be estimated. As an initial step, table values for absorption and scattering from several different sources were used. The full table of material parameters is given in Appendix A.

Whereas there are many table values for absorption coefficients, scattering coefficients are more difficult to measure and may vary depending on the software used. In ODEON, scattering coefficients for 707 Hz are provided by the user, which are then extrapolated to the full frequency range using a built-in algorithm. For the simulations in this project, table values were primarily used. When these were unavailable, mid-frequency scattering was estimated from the characteristic depth of the structure in question, as discussed in [51].

To improve the quality of the absorption coefficient estimates, calibration by comparison to modern reverberation time measurements (measured by the integrated impulse response method [52] and a B&K type 2270 analyzer) was performed. As the modern interior of the church differs so significantly from the historical configuration (see, e.g., Figures 1 and 3–5), only a few materials can be calibrated: the vaulted ceilings, the glass windows, and the stone floor. These are assumed to have the same material properties as during the 1470s, whereas the walls which were historically covered by plaster are now bare and can not be calibrated.

The calibration is performed using ODEON's built-in genetic algorithm optimization tool, in a digital model corresponding to the modern church. This model comprises the exterior shell, modern wooden pews, and the main altar. This corresponds to the conditions under which the modern measurements were performed. The average error of the simulated reverberation time in the calibration model compared to measurements is within ±1.4 JND (see in Figure 8 for detailed comparison for each octave band and position). The calibrated material parameters were then used in the historical digital model.

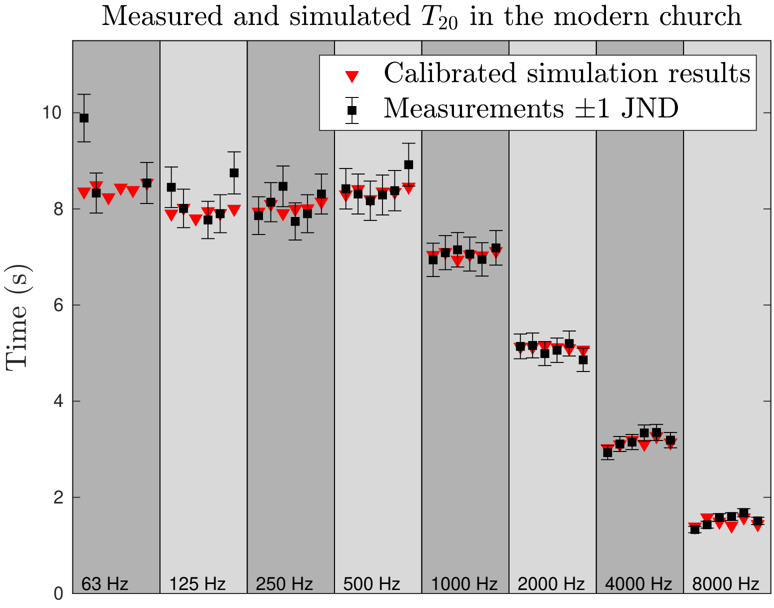


Figure 8. Comparison between simulated and measured T_{20} values in the modern church. The red triangles show the results of the calibrated simulation, and the black boxes correspond to measurements together with frequency-dependent 1 JND error bars. The results for six different listener–receiver combinations are shown across center octave-band frequencies from 63 to 8000 Hz. For low frequencies, some values could not be extracted from the measurements and only the simulated results are shown. The average error in JND is highest in the 63 Hz and 8000 Hz octave bands with 1.2 and 1.4, respectively; the remaining octave bands have values between 0.2 and 0.7.

3.3.3. Listener and Source Positions

Sound source and sound receiver positions were chosen according to the goals of the simulations. The source positions for room acoustic simulations are shown in Figure 9. The nuns' gallery was identified as a location of interest, and source P1 and a receiver are located there. Similarly, both a source (P2) and a receiver are in the monks' choir. In addition, one sound source was located close to the high altar (point P3). Additional receiver positions are distributed on the floor of the nave. In the auralization, twelve nuns and eight monks modeled as point sources are placed in their respective choir stalls in the gallery and choir.

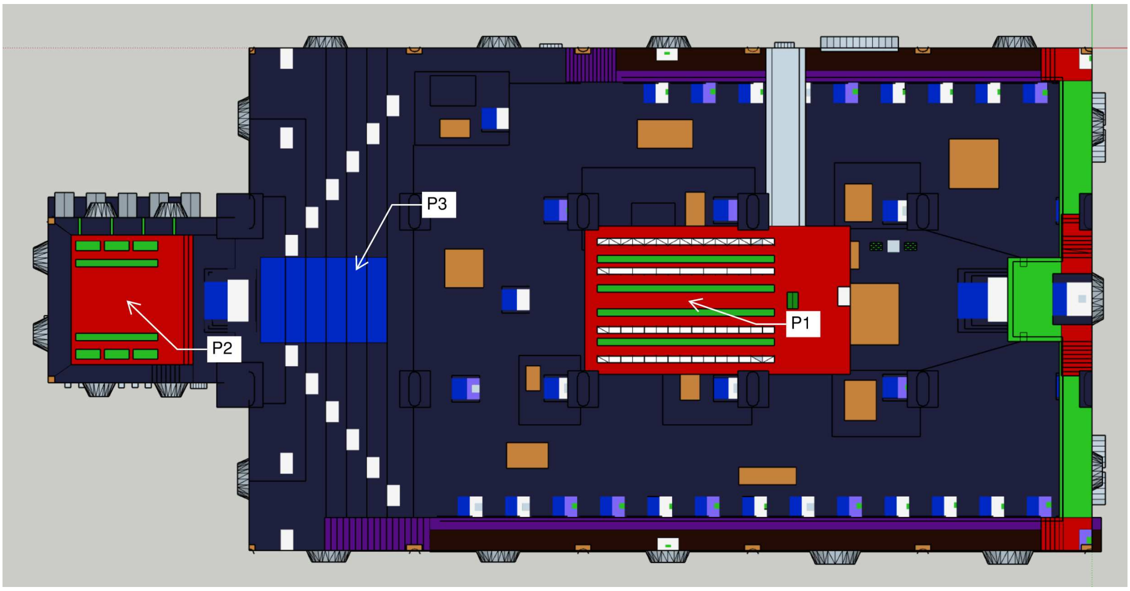


Figure 9. Top-down view of the digital reconstruction of the abbey. North is up in the figure. The monks' choir is the red area to the left, and the nuns' gallery is the red area in the center. Three source positions—P1, P2, and P3—are marked.

4. Room Acoustic Analysis

The sound field in the reconstructed church was evaluated from room acoustic simulations. Three questions were specifically examined. First, three different visitor conditions were evaluated. The three conditions corresponded to an empty church, an ordinary sext (about 30 individuals), and a more festive event (about 140 individuals). The sext condition corresponds to the situation targeted in the auralization, with twelve nuns in the gallery, eight monks in the choir, and about ten other visitors in the central-eastern part of the nave. In the festive condition, denoted "full", there were 70 nuns in the gallery, 13 monks in the choir, and about 70 visitors in the nave.

Second, the sound fields within the choir and gallery are examined. These locations are important for the religious practices of the order, and acoustic analysis of these spaces may provide new insights. Third, the sound field in the nave is examined.

The auralizations are provided as supplementary material, and are not further analyzed in this text.

Four room acoustical parameters are presented, as defined in ISO 3382-1:2009 [52]. The reverberation time T_{20} is presented, due to its traditional importance in acoustics. It is primarily useful as a tool for comparison to other spaces, as the reverberation time is a commonly measured acoustic characteristic. The early decay time (EDT) is also presented. It describes the rate of sound energy decay in the first parts of the impulse response, and is closely related to perceived reverberation and the presentation quality of Gregorian chant.

According to guidelines proposed by Martellotta et al. [4], EDT values in the range 2.1 s to 4.2 s are appropriate for churches. In addition, C_{80} is presented as an indicator of the clarity and intelligibility of chant. As the liturgical practices of this region and time do not contain significant spoken elements, C_{50} (speech clarity) is not presented separately. For concert halls, C_{80} values of above 0 dB are usually considered "good", but no such guidelines have been defined for worship spaces. Finally, the sound strength (G) is presented. This value shows the total sound pressure level (SPL) at the receiver, as compared to what would be perceived from the sound source in free field at a distance of 10 m. G is positive in enclosed spaces.

A summary of the results is shown in Table 1. This table gives values for T_{20}, C_{80}, EDT, and G, averaged according to ISO 3382-1:2009 [52]. In addition, the results are divided into categories based on the location of sources and receivers. If the source and the receiver are both within the gallery (point P1 in Figure 9), the data is denoted "Gallery". If both source and receiver are in the choir (point P2 in Figure 9), the data is denoted "Choir". Finally, if the sound has travelled through the nave, the data is categorized as "Nave". This includes when either the source or the receiver is located in the nave, but also the conditions where one is in the choir and one is in the gallery.

Table 1. Room acoustic parameters in various configurations of the reconstructed abbey. The values have been averaged according to ISO 3382-1:2009 [52].

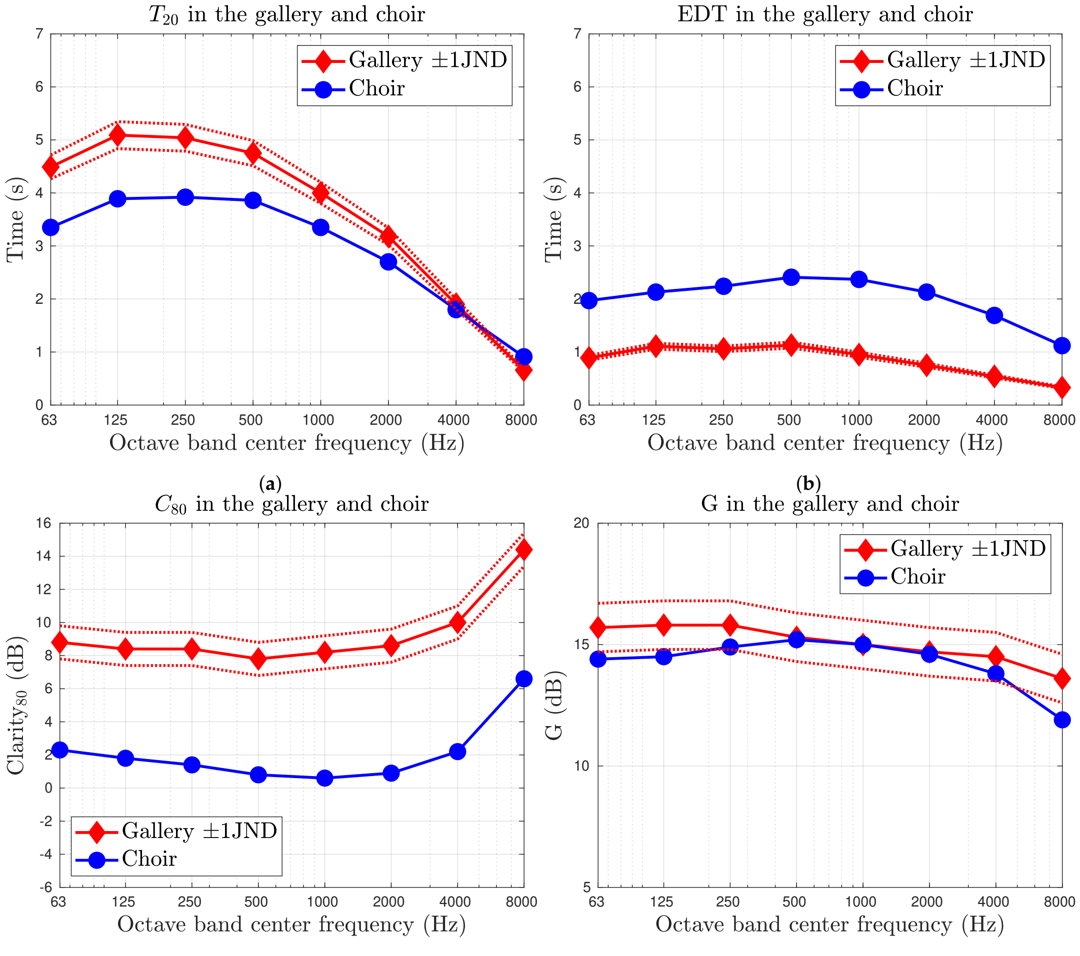
	Gallery			Choir			Nave		
	Empty	**Sext**	**Full**	**Empty**	**Sext**	**Full**	**Empty**	**Sext**	**Full**
T_{20} (s)	5.17	4.90	4.00	4.06	3.89	3.26	5.37	5.19	4.36
EDT (s)	1.21	1.10	0.62	2.41	2.33	2.21	5.47	5.29	4.55
C_{80} (dB)	8	8	10	1	1	1	−10	−10	−10
G (dB)	16	16	15	15	15	15	6	6	4

One result seen in Table 1 is that the number of visitors and the visitor condition have an impact on all room acoustic parameters presented. There are no or almost no differences between the empty and the sext condition, but the full condition leads to significant decreases of EDT and T_{20}. This is an expected result of the increased amount of absorption when the number of visitors increase. Increased absorption should also lead to an increased C_{80} and a decrease in G. These patterns can be seen, but are weak. The only differences larger than 1 JND are seen for C_{80} in the gallery and G in the nave.

The results in Table 1 show a clear difference between the sound field within the gallery and within the choir, compared to the other configurations. All of the room acoustic parameters presented above are significantly different for these two subspaces as compared to the rest of the church. The subspaces are analyzed separately in Section 4.1. The results for the rest of the church is presented in Section 4.2.

4.1. Acoustics within the Choir and Gallery

The acoustic simulation results within the gallery (source and receiver both near point P1 in Figure 9, reflecting the experiences of the nuns), and within the choir (source and receiver near point P2, reflecting the experience of the monks) are presented in Figure 10. In these graphs, room acoustical parameters T_{20}, EDT, C_{80}, and G are shown in octave band resolution. The gallery data are shown together with lines denoting ±1 JND, such that when the choir data falls within these lines there is no perceivable difference between the two locations.



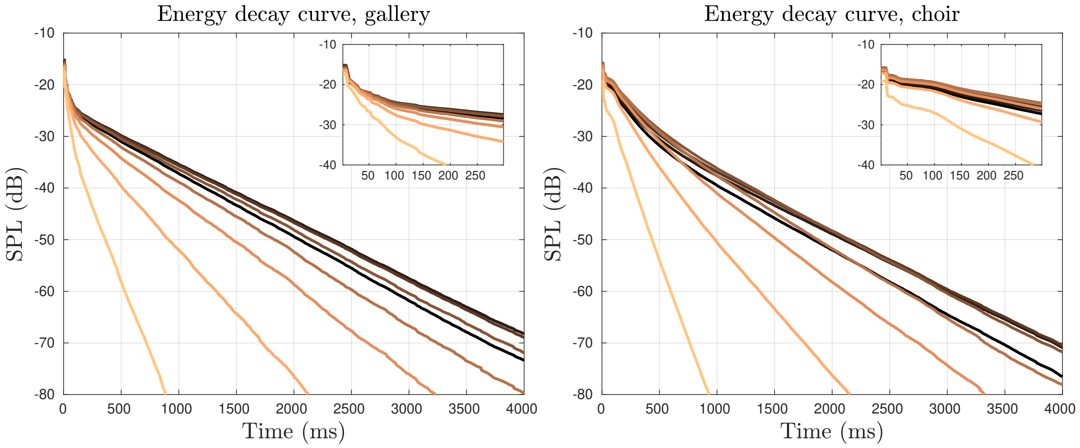
(**a**) (**b**)

(**c**) (**d**)

Figure 10. Graphs showing (**a**) T_{20}, (**b**) early decay time (EDT), (**c**) C_{80}, and (**d**) G in octave band resolution in the reconstructed abbey choir and gallery. The results for the gallery (red) are shown with dotted lines corresponding to ±1 JND. When the results for the choir (blue) falls within these lines, there is no perceivable difference between the two data sets.

The reverberation time is significantly lower in the choir compared to the gallery. A review of the full results in Table 1 shows that the results in the gallery and the results in the rest of the church only differ by ±1-2 JND, whereas the discrepancy with the choir is larger.

The second reverberation parameter, EDT, shows a different pattern. The EDT is much smaller than T_{20} for both positions, and the EDT in the gallery is significantly lower than that in the choir. The large discrepancy between the EDT and the T_{20} may indicate that the decay curves in these spaces follow a multi-slope decay pattern, which would be consistent with these spaces acting as acoustically distinct subspaces. The energy decay curves are more closely investigated in Figure 11.



(**a**) (**b**)

Figure 11. Energy decay curves in the (**a**) nuns' gallery and (**b**) monks' choir in the reconstructed abbey. The small insets show the early parts of the decay. Each line corresponds to the results in one octave band.

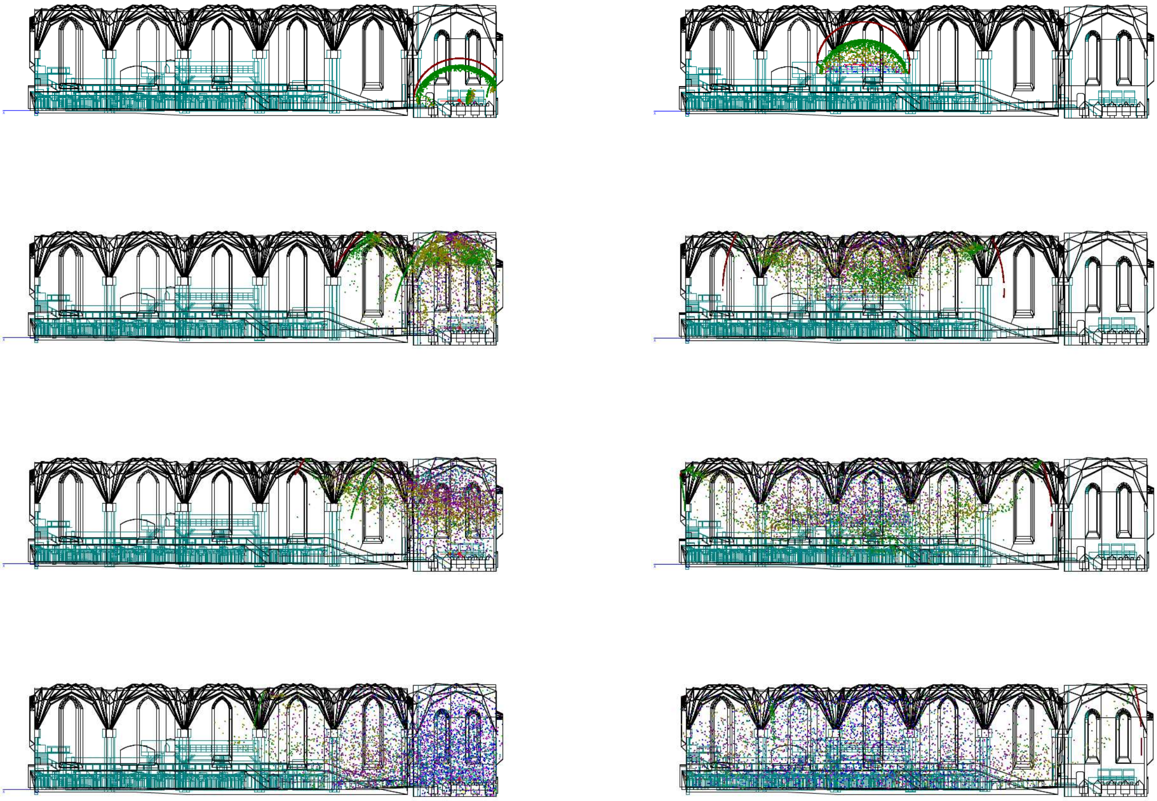
Reviewing the energy decay curves in Figure 11 shows a distinctly curved decay pattern in both gallery and choir. As expected from the results in Figure 10, the acoustic energy in the gallery decreases rapidly in the early parts of the decay, and then flattens out to a much slower decay rate. This is consistent with the formation of an acoustic subspace within the gallery, coupled to the larger, more reverberant space in the nave. As the acoustic energy within the subspace decreases, the reverberance of the nave becomes more dominant. The transition seems to occur after about 100 ms. The energy decay curve in the choir seems more complex, and is likely influenced by the three close, hard walls on three of four sides as well as the more distant vaulted ceiling. It is possible that these decay curves are formed by more than two distinct decay patterns. It can, however, be seen that the decay process after about 750 ms is more smooth. This transition time is later than in the gallery.

Returning to the results in Figure 10, it is seen that C_{80} is also significantly different between the gallery and the choir across all octave bands. Clarity is much better in the gallery compared to the choir, which is in turn much better than what is found in the rest of the church (Table 1). The reason for these differences can be explained by the physical configuration and particularly the effects of the ceiling vaults. A time series illustrating a raytracing model of the energy distribution over time is shown in Figure 12.

On the left hand side in Figure 12, the energy distribution over time for a sound source located in the choir is shown. The energy emitted from the source travels upward as time progresses, to be reflected from the ceiling vaults. By 80 ms, the third figure from the top, the reflected wavefront can be seen in the space above the choir. Consequently, it has not reached the receiver by 80 ms, and thus has a detrimental effect on C_{80}.

On the right hand side in Figure 12, the corresponding time series for a source and receiver in the gallery is shown. As the gallery is closer to the ceiling, the strong reflected wavefront from the ceiling has already reached the receiver in the gallery by 80 ms, thus improving C_{80}.

The G within the gallery and choir can be seen in Figure 10d. As shown in the graph, there are no or very small differences between the gallery and the choir. The results in both spaces are, however, significantly better than outside these spaces. This can be explained by strong reflections, from the ceiling in the gallery and by the walls in the choir. The positions of these surfaces ensure that a significant amount of sound energy is reflected back to the space, thus increasing the total G.



(**a**) (**b**)

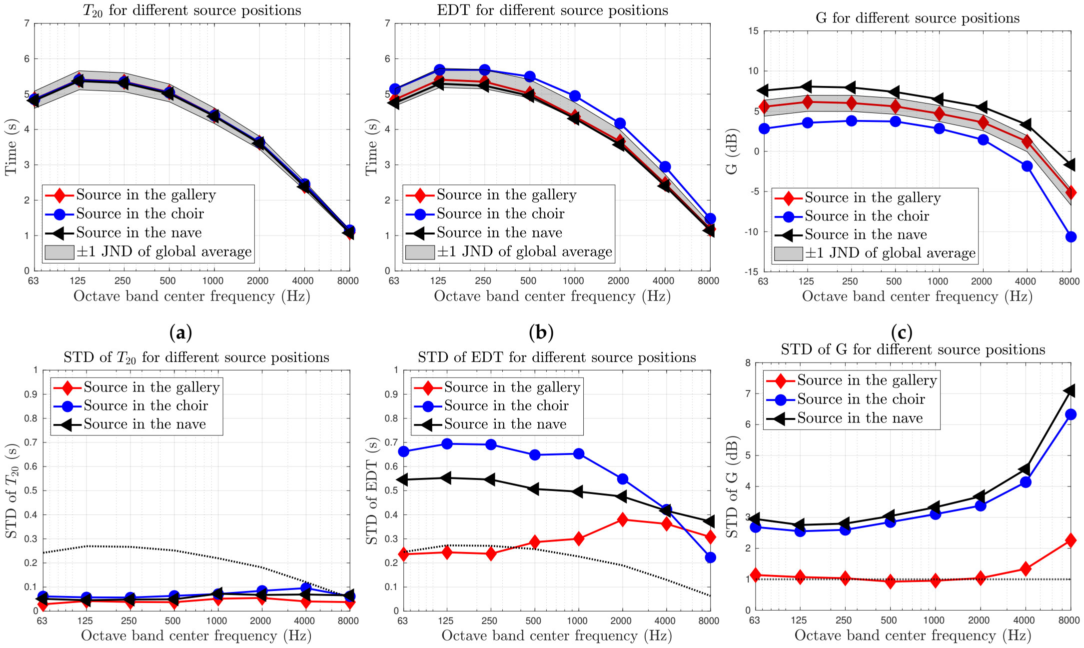
Figure 12. Images illustrating the dispersion of acoustic energy, as approximated by an acoustic particle model, over time. Subfigure (**a**) shows the progression when the source is located in the choir, and subfigure (**b**) shows when the source is located in the gallery. From top to bottom, the pictures show snapshots at 20 ms, 60 ms, 80 ms, and 120 ms.

Finally, some brief comments are made regarding sound transmission between the gallery and choir, corresponding to the acoustic perception of monks from the nuns' position and vice versa. It is found that acoustic transmission between the two locations is very similar to the transmission between choir and nave. As those results are presented in Section 4.2, no further comments are made here.

4.2. Acoustics in the Nave

Within the nave, the acoustic simulations aim first at characterizing the acoustics of the space itself and, second, at evaluating any difference between various source locations. As such, the data in this section are presented separately, according to the source position. The three source positions are in the gallery (point P1 in Figure 9, nuns' position), in the choir (point P2, monks' position), or in the nave, by the high altar (point P3). T_{20}, EDT, and G results are shown in Figure 13. As the results in this section are in general spatially averaged, the standard deviations are also shown as an estimate of the spatial variations between listener positions.

The average reverberation time in the reconstructed abbey is shown in octave band resolution Figure 13a. Results for the three different source positions are shown, and are very similar; also shown in the graph is a gray area delimiting values that are within 1 JND of the global average. All lines fall entirely within this area, indicating that there are no significant differences in the reverberation time in the nave depending on the location of the sound source; also shown in Figure 13d are the standard deviations of T_{20} for each of the source positions. These are shown together with the global JND for reverberation time. All standard deviations fall below the JND line, indicating that the variations of the reverberation time are on average imperceptible. This indicates that the late acoustic response within the nave was, in general, diffuse and not characterized by significant spatial variations.



(**a**) (**b**) (**c**)

(**d**) (**e**) (**f**)

Figure 13. Overview of room acoustic parameters in the reconstructed church. Results are shown for three different source positions in octave band resolution. (**a**) The average T_{20} and (**d**) its standard deviation. (**b**) The EDT and (**e**) its standard deviation. (**c**) The G and (**f**) its standard deviation. All standard deviations are shown compared to the global average JND of that parameter (dotted line). The gray areas in graphs (**a**–**c**) delimit values within 1 JND of the global average.

In general, the reverberation time is long, but comparable to similar spaces. T_{20} is significantly decreased compared to modern day measurements, as expected by the discussion in Section 3.1.1. In Table 2, the mid-frequency reverberation time (T_m) is shown together with measurement results from the modern day church (from in [43]) and some other Gothic churches around the Baltic Sea (from in [11]). These results allow for a basic comparison between the reconstructed abbey and similar spaces.

Table 2. Mid-frequency reverberation times for some Gothic churches in Northern Europe.

Church	**Volume (m^3)**	**T_m (s)**
Swedish church, current configuration	29,000	7.79
Swedish church, historical configuration	29,000	5.15
Church of our Lady, Krakow, Poland	9500	6.5
Church of St Thomas, Lipsk, Germany	18,000	4.05
Marien Church, Lübeck, Germany	100,000	5.50

Spatially averaged EDT values for each source position are shown in Figure 13b. Again, the gray area delimits values within 1 JND of the global average. The variations for different source positions are greater than those seen for T_{20}, but only the results for the choir are significant compared to the JND. This difference is, however, small. Comparing the global EDT values to the guidelines proposed by Martellotta et al [34] shows that the average EDT exceeds the recommended values for the reconstruction.

In Figure 13e, the standard deviation of the EDT is shown for each source position together with the global JND. The standard deviation for a source in the gallery is smaller than the others, and falls below 1 JND for low frequencies. The typical variations for the sources in the nave and in the choir exceed 1 JND. This shows that the perceived reverberance of sources in those two locations varies significantly depending on the listener's position. It also indicates that the small difference in average EDT between sources in the choir and elsewhere (seen in Figure 13e) may be too small to be relevant, compared to the variations between receiver positions.

G in the nave is shown in Figure 13e, in octave band resolution for each of the three source positions under consideration. The gray area delimits values within 1 JND of the global average. In this case, there are significant differences between the average results for the three positions. Sound emitted from the source by the high altar is on average heard louder, and sound from sources in the choir are on average heard more quietly. Part of the explanation for this fact may be that there are more receivers with a direct line-of-sight to the sound source in the nave than for the sources in the gallery and the choir. The contribution of the direct sound has a significant impact on the overall sound pressure level, and thus on G.

The standard deviations of G is also shown, in Figure 13f. Again, these are shown together with the corresponding JND. This graph shows that for sources located in the choir or by the high altar, the spatial variations of G are of a very similar magnitude, and significantly larger than when the source is located in the gallery. This can partially be explained by the position of the gallery in the middle of the church. G is affected by the distance between the source and the receiver, and the gallery's position ensures that the distance between source and receiver varies minimally.

C_{80} simulation results are shown in octave band resolution in Figure 14. As C_{80} is not expected to be uniform in a space, all measurement points are shown, together with a line indicating the spatial average. The average C_{80} for sources in the choir and gallery fall in the region of −15 dB to −5 dB, significantly below what would be characterized as "good" for a concert hall. Clarity for sources located by the high altar is better, at about −5 dB to about 2 dB.

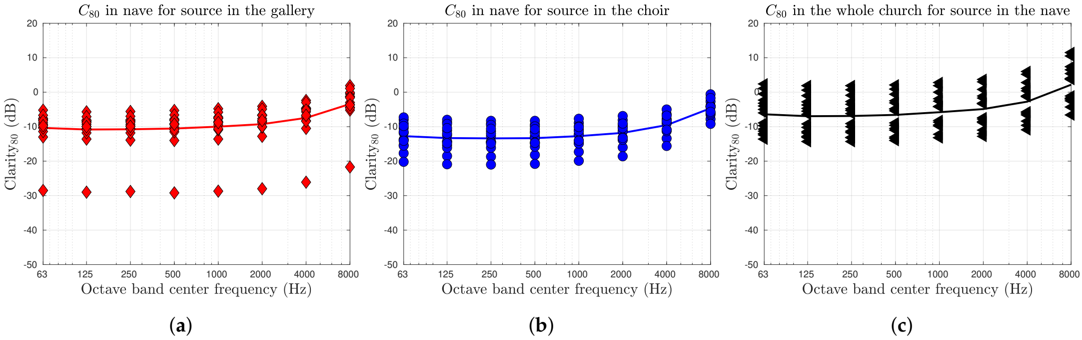
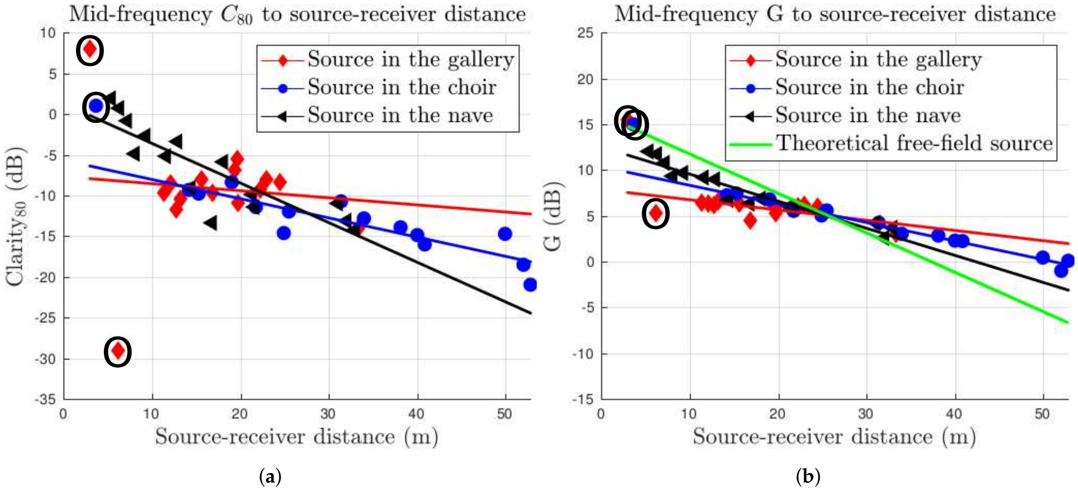


Figure 14. C_{80} in octave bands simulated for the recreated church, shown for 13 different receiver positions. The spatial mean (line) is also shown. (**a**) C_{80} when the sound source is located in the nuns' gallery. One position produces significantly lower clarity. (**b**) C_{80} when the sound source is located in the choir. (**c**) C_{80} when the sound source is located by the high altar in the nave.

Although C_{80} is not expected to be uniform in a space, the results in Figure 14 show that most values fall in clusters around the average value. However, there is one outlier in Figure 14a, when the source is located in the gallery. This estimate comes from the receiver located underneath the gallery, below point P1 in Figure 9. It is separated from the sound source by the physical structure of the gallery itself, and there is no direct line of sight. A review of the acoustic simulation reveals that the first sound energy reaching this location from the source in the gallery is a second-order reflection, showing that sound energy reaching this location from the gallery consists solely of reflections of order two or higher. This significantly decreases the cohesion and energy level in early parts of the impulse response, leading to a very poor clarity of sources in the gallery as perceived by listeners beneath it.

The acoustic field in the nave is further analyzed to determine whether the spatial variations for sources in the choir and by the high altar, seen in Figures 13e,f and 14, are caused by acoustic subspaces. It is found that the spatial variations can be explained well by the distance between source and receiver. Linear regression models for mid-frequency G and C_{80} to source–receiver distance are shown in Figure 15, and the R^2-values are presented in the caption. Except for sources in the gallery, more than 70% of the typical spatial variations in G and C_{80} can be explained by the source–receiver distance.

In Figure 15, the results from within the gallery, within the choir, and the outlier found in Figure 14 are marked, and not included in the linear regression models. They deviate significantly from the pattern defined by the regression line, indicating that the assumption that these locations are governed by the acoustic properties of a certain subspace is accurate. However, no other measurement points show a similar deviation from the prediction. Thus, there are no indications that there are distinct acoustical subspaces in the reconstructed church except those already identified.



(**a**) (**b**)

Figure 15. (**a**) Mid-frequency C_{80} and (**b**) mid-frequency G as they vary with source–receiver distance; also shown are linear regression models estimating the influence of source–receiver distance for each source locations. R^2 values for the models in subfigure (**a**): Source in the gallery: ($R^2 \approx 0.06$), choir: ($R^2 \approx 0.78$), and nave ($R^2 \approx 0.75$). In subfigure (**b**), the corresponding values are gallery: $R^2 \approx 0.48$, choir: $R^2 \approx 0.98$, and nave: $R^2 \approx 0.92$. In subfigure (**b**), the regression line for a theoretical omnidirectional source in free field is also shown. Highlighted data points have been excluded from the regression lines. These correspond to configurations entirely within the gallery or choir and the outlier case when the source is in the gallery and the receiver below it.

Another pattern emerges in the evaluation of source–distance dependence, showing that the influence of distance varies depending on the location of the sound source. As shown in Figure 15a, C_{80} decreases as source–receiver distance increases and the reverberant field becomes more dominant. The linear regression quantifies the influence of distance for each source position. The effects of distance is most strongly seen when the sound source is located in the nave, by the high altar. Moving the sound source to the choir decreases this effect slightly. When the sound source is in the gallery, the effect is almost entirely gone, and clarity is not significantly affected by increasing the distance to a sound source in the gallery. This analysis is confirmed by a review of the R^2-values for the respective linear regression models. These indicate that only about 6% of the variations of C_{80} in the nave for sources in the gallery can be explained by the distance between receiver and source. The values are significantly higher for sources in the choir and nave.

In Figure 15b, the distance dependence of G is examined and compared to the SPL decay of a free-field source. To facilitate comparison, the free-field condition is approximated with a linear regression model and is not normalized to the source's SPL at a distance of 10 m. As would be expected, the G decreases with increased source–receiver distance across all data series. However, the slope of the decays within the church are softer than that for the free field condition. This is due to the sustained reflections of the enclosed space. The G decay is fastest for the free field condition, followed by the source by the high altar. For sources in the choir, the effect is smaller and for sources in the gallery smaller still.

These results indicate that the acoustic perception of sources by the high altar (P3 in Figure 9) strongly depends on the listener's position, which implies that the sound field created by such a source is less homogeneous. This may be explained by the lack of nearby reflecting structures around that source position, as compared to the choir and gallery. Accordingly, there are no strong early reflections supporting the transmission of sound from this location. Thus, acoustic energy is dispersed in all directions, whereas the vaulted ceiling above the gallery and the walls of the apse reinforce early reflection and direct the the spread of acoustic energy from the gallery and the choir into the nave.

There is also a significant difference between the choir and the gallery in distance dependence, which can be understood further by again turning to the raytracing model shown in Figure 12. From these graphs, it appears that, after about 100 ms, the acoustic energy is distributed much more evenly in the nave when the source is located in the gallery, rather than the choir. This behavior can be explained by two factors. First, the gallery is located in a much more advantageous location in the middle of the church. This leads to a much more even distribution of sound energy initially. Second, the gallery is larger and interacts with multiple ceiling vaults, while the acoustic energy from the choir only is reflected by the vaults of one bay. The reflections from multiple ceiling vaults leads to a much greater scattering of the acoustic field, and a greater diffusion overall. These two factors together give a more even distribution of acoustic energy when the source is located in the gallery.

5. Discussion

The room acoustic analysis of the reconstructed abbey reveals the presence of two distinct acoustical subspaces, which coincides with locations of significant liturgical importance. The nuns' gallery and the monks' choir are characterized by shorter EDT, greater G, and improved C_{80} compared to the rest of the church. These better acoustic conditions facilitate the auditory elements of liturgical practices, which are a fundamental part of the monastical congregations' daily tasks. Examples of such subspaces have been found in many other worship spaces [31–33] and are sometimes referred to as a "church within a church", indicating their role as an exclusive environment for the initiated. The presence of two such locations, rather than one, is an expected consequence of the presence of two separate monastic enclosures within the same church.

Despite their similarity as acoustic subspaces with improved conditions for liturgical activities, there are distinct differences between the gallery and the choir. The proximity of the gallery to the vaulted ceiling both improves the clarity within it and results in a rather homogenous sound field in the nave when a sound source is located in the gallery. Within the choir, as compared to within the gallery, the reverberation time is more significantly different to that in the nave. This could imply a weaker coupling between the acoustic subspace in the choir and the nave, than between the gallery and the nave. This could be a reason for the poor acoustic transmission from the choir to the nave.

Such differences may reflect a religious intent, irrespective of whether there was an acoustical intent in the design. The position of the gallery reflects a religious intent to premier the nuns within this abbey, and the improved acoustic transmission from this location supports this intent. Its location makes the gallery both acoustically and visually characteristic for the whole church, and establishes the nuns' position as central within the monastic order. The reflections from the ceiling vaults cause the nuns' chants to be perceived as coming "from above". This could reflect an acoustic intent of making the nuns sound more heavenly.

The spatial variations in the church can be heard in the auralizations (as presented in the supplementary material) of elements from a Friday sext. When both source and listener is in the gallery, or both in the choir, intelligibility is acceptable and the acoustics support chants, prayers, and responsories sufficiently to seem plausible. Reverberation from the nave can be heard, but is not strong enough to dominate the sound field. The services as perceived outside the respective monastical subspace give a very different impression. Individual syllables can not be distinguished, the locations of the sound sources are difficult to determine and the reverberance dominates perceived sound. This is the case for the perception of monks from within the gallery, nuns from within the choir and both from within the nave. Monks and nuns sound distant, yet omnipresent. Liturgical interactions between monks, nuns or people in the nave were thus likely not possible during daily services.

Further analysis of the auralizations themselves, with listening tests, could lead to additional insights regarding the experiences of historical visitors to this place. For example, the introduction of HRTFs is needed in order to evaluate the perceptual impact of the nuns' elevated position, to further evaluate the theory of their voices sounding more heavenly. Such research could show more clearly the perceptual differences between the experiences of the monastical congregations and the pilgrims. Furthermore, it may be possible to use the results of the simulation as a tool for VR performances, where singers experience the "live" simulated acoustics of the historical space as they sing. Such setups may minimize the effects of recording in an anechoic chamber.

The reverberation time within the nave is long but comparable to other churches around the Baltic sea. Although no comparison could be made for other Scandinavian or Nordic churches, comparisons to Gothic churches in other countries around the Baltic Sea (Germany and Poland) indicate some similarities in the acoustic cultural heritage. However, as the sample size is so small, further research is needed before any conclusions can be drawn.

The significant reverberation within the nave, and especially the long EDT, indicates that the space in the sext configuration is not suitable for the Gregorian chants usually performed there. This conclusion is supported by listening to the auralizations. However, in the more festive condition examined, with an increased number of visitors and members of the congregations, the EDT is decreased sufficiently within the nave to be within Martellotta et al.'s [34] guidelines for Gregorian chant. While there are no indications that this was intentional, it implies that during events aimed at a more general public, the acoustics of the church supported such events. Accordingly, the acoustics could be considered somewhat self-regulatory; when there are few listeners in the nave, acoustics are sufficient only for the important monastical congregations, and when the number of visitors in the nave increases, the larger number of visitors experience an acceptable acoustic field.

However, caution should be applied when making quantitative comparisons of archaeoacoustical simulations to guidelines or measurements, such as the comparisons made above. The rate of uncertainty in the model may be significant, both due to uncertainties in the material parameters and in the geometric modeling techniques [53,54]. The calibration procedure employed in this paper reduces these uncertainties, but can not remove them entirely.

Despite these words of caution, it should be noted that the major conclusions of this work concern the acoustic subspaces in the gallery and nave. These effects are primarily caused by the geometrical configuration of the space, which has been established based on the thorough historical research presented in Section 3.1. As such, the qualitative conclusions regarding improved acoustic conditions for the monastical congregations holds, despite the uncertainties presented regarding the absolute values of calculated parameters.

6. Conclusions

The room acoustic analysis in this paper shows that acoustical subspaces were formed in locations of religious and liturgical importance. These acoustical subspaces offered improved acoustics for the monastical congregations for which the church was built. Of the two congregations, the nuns were more central, and this is reflected in the design and acoustics of the church.

Although some of the results of the acoustic analysis of the gallery may have been possible with a less accurate digital model, the thorough acoustical analysis has benefited from the substantial historic research underlying it and finds its validation from it. The precise determination of the gallery's size, elevation, position, and form has been shown to have consequences for the sound field within the gallery and within the nave. This shows that archeo-acoustical modeling benefits from tight collaboration between acousticians, historians, and 3D-artists.

Supplementary Materials: The following are available online at https://www.mdpi.com/2076-3417/11/4/1586/s1.

Author Contributions: Conceptualization, H.A., M.B., C.A., D.B.H., E.L.S. and K.S.L.; methodology, M.B., C.A., E.L.S. and K.S.L.; formal analysis, H.A. and M.B.; investigation, M.B., C.A., E.L.S. and K.S.L.; writing—original draft preparation, H.A. and C.A.; writing—review and editing, H.A. M.B., C.A. and D.B.H.; visualization, H.A. and C.A.; supervision, D.B.H. and E.L.S.; funding acquisition, D.B.H., E.L S. and K.S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Swedish Research Council grant number 2016-01784.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding authors.

Acknowledgments: The authors express gratitude to: Brekke & Strand Akustik AB for providing ODEON 16.0 and Cubase 11.0 licences; Marcin Brycki from Brekke & Strand Akustik AB for leading the recording sessions and performing the mixing; Milab Microphones AB for providing microphones for recording; the singers involved in the recordings; Nikolaos-Georgios Vardaxis and Erling Nilsson from Engineering Acoustics LTH for valuable inputs; Tim Näsling from Brekke & Strand Akustik AB for useful tips and discussions during room acoustic modelling; Stefan Lindgren at LU Humanities Lab for the 3D scan of the church; Mattias Hallgren at Traditionsbärarna for input regarding historical construction.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. Material Parameters for Acoustic Simulation

Table A1. Material coefficients used for the room acoustic simulations and auralization of the reconstructed abbey.

Material	**Absorption Factor**								**Mid-Frequency Scattering**
	63 Hz	**125 Hz**	**250 Hz**	**500 Hz**	**1000 Hz**	**2000 Hz**	**4000 Hz**	**8000 Hz**	
Thin silk textile, freely suspended [1]	0.05	0.05	0.06	0.39	0.63	0.70	0.73	0.73	0.01
Wool textile [1]	0.07	0.07	0.31	0.49	0.75	0.70	0.60	0.60	0.01
Thick wool (carpet) [1]	0.02	0.02	0.06	0.14	0.37	0.60	0.65	0.65	0.01
Heavy velvet [1]	0.03	0.03	0.03	0.15	0.4	0.50	0.50	0.50	0.01
Thick linen against stone [1]	0.01	0.01	0.02	0.05	0.15	0.30	0.40	0.40	0.01
Wooden construction, not painted [2]	0.09	0.09	0.09	0.08	0.08	0.10	0.07	0.07	0.60
Wooden construction, painted [2]	0.11	0.11	0.11	0.10	0.10	0.10	0.07	0.07	0.40
Wooden decoration, painted [2,3]	0.12	0.12	0.12	0.15	0.15	0.18	0.18	0.19	0.99
Hollow wooden structure, painted [1]	0.40	0.40	0.30	0.20	0.17	0.15	0.10	0.10	0.60
Plastered brick [6]	0.102	0.029	0.144	0.097	0.007	0.016	0.008	0.003	0.20
Limestone [6]	0.028	0.074	0.005	0.034	0.09	0.028	0.084	0.009	0.003
Plastered limestone [4]	0.07	0.06	0.05	0.05	0.04	0.04	0.04	0.04	0.001
Ceiling vaults [5]	0.12	0.09	0.09	0.05	0.04	0.03	0.03	0.03	0.30
Leaded glass windows [6]	0.254	0.259	0.24	0.016	0.101	0.039	0.495	0.003	0.14
Iron lattice [5]	0.01	0.01	0.01	0.02	0.06	0.03	0.03	0.03	0.001
People [1]	0.62	0.62	0.72	0.80	0.83	0.84	0.85	0.85	

[1] From an ODEON standard material [50]; [2] Suarez et al. [55]; [3] Alonso et al. [56]; [4] Postma et al. [51]; [5] Own data; [6] Determined by genetic algorithm optimization, see Section 3.3.2.

References

1. UNESCO. *Concept of Digital Heritage*; UNESCO: London, UK, 2019.
2. UNESCO. *Convention for the Safeguarding of the Intangible Cultural Heritage*; UNESCO: London, UK, 2003.
3. Lubman, D.; Kiser, B.H. The History of Western Civilization Told through the Acoustics of its Worship Spaces. In Proceedings of the 17th International Congress on Acoustics, Rome, Italy, 2–7 September 2001.
4. Martellotta, F. Optimal Reverberation Conditions in Churches. In Proceedings of the International Congress on Acoustics 2007, Madrid, Spain, 2–7 September 2007; Sociedad Espanola de Acustica: Madrid, Spain, 2007.
5. Howard, D. Recordings of Music Written for St. Mark's. In *Word, Image, and Song: Essays on Early Modern Italy*; University of Rochester Press: Rochester, NY, USA, 2013; Volume 1, pp. 89–100.
6. Navarro, J.; Sendra, J.J.; Muñoz, S. The Western Latin church as a place for music and preaching: An acoustic assessment. *Appl. Acoust.* **2009**, *70*, 781–789. [CrossRef]
7. Pedrero, A.; Ruiz, R.; Díaz-Chyla, A.; Díaz, C. Acoustical study of Toledo Cathedral according to its liturgical uses. *Appl. Acoust.* **2014**, *85*, 23–33. [CrossRef]
8. Aletta, F.; Kang, J. Historical Acoustics: Relationships between People and Sound over Time. *Acoustics* **2020**, *2*, 128–130. [CrossRef]
9. Đorđević, Z.; Novković, D.; Andrić, U. Archaeoacoustic Examination of Lazarica Church. *Acoustics* **2019**, *1*, 423–438. doi:10.3390/acoustics1020024. [CrossRef]
10. Elicio, L.; Martellotta, F. Acoustics as a cultural heritage: The case of Orthodox churches and of the "Russian church" in Bari. *J. Cult. Herit.* **2015**, *16*, 912–917. [CrossRef]
11. Girón, S.; Álvarez-Morales, L.; Zamarreño, T. Church acoustics: A state-of-the-art review after several decades of research. *J. Sound Vib.* **2017**, *411*, 378–408. [CrossRef]
12. Kosała, K. Calculation Models for Acoustic Analysis of St. Elizabeth of Hungary Church in Jaworzno Szczakowa. *Arch. Acoust.* **2016**, *41*, 485–498. [CrossRef]
13. Niemas, M.; Sadowski, J.; Engel, Z. Acoustic Issues of Sacral Structures. *Arch. Acoust.* **1998**, *23*, 87–104.
14. Kanev, N. Resonant Vessels in Russian Churches and Their Study in a Concert Hall. *Acoustics* **2020**, *2*, 399–415. doi:10.3390/acoustics2020023. [CrossRef]
15. Swedish Research Council. The Multisensory World of Vadstena Abbey in the Late Middle Ages. 2016. Available online: https://www.swecris.se/betasearch/details/project/201601784VR?lang=en (accessed on 11 January 2021).
16. Scarre, C.; Lawson, G. *Archaeoacoustics*; McDonald Institute for Archaeological Research: Cambridge, UK, 2006.
17. Boren, B.; Longair, M. A Method for Acoustic Modeling of Past Soundscapes. In Proceedings of the Acoustics of Ancient Theatres Conference, Patras, Greece, 18–21 September 2011; European Acoustics Association and the Hellenic Institute of Acoustics: Patras, Greece, 2011.
18. Berardi, U.; Iannace, G.; Maffei, L. Virtual reconstruction of the historical acoustics of the Odeon of Pompeii. *J. Cult. Herit.* **2016**. [CrossRef]

19. Martellotta, F.; Álvarez-Morales, L. Virtual acoustic reconstruction of the church of Gesú in Rome: A comparison between different design options. In Proceedings of the Forum Acusticum 2014, Krakow, Poland, 7–12 September 2014; The Polish Acoustical Society: Krakow, Poland, 2014. [CrossRef]
20. Alonso, A.; Sendra, J.J.; Suárez, R.; Zamarreño, T. Acoustic evaluation of the cathedral of Seville as a concert hall and proposals for improving the acoustic quality perceived by listeners. *J. Build. Perform. Simul.* **2014**, *7*, 360–378. [CrossRef]
21. Sender, M.; Planells, A.; Perelló, R.; Segura, J.; Giménez, A. Virtual acoustic reconstruction of a lost church: application to an Order of Saint Jerome monastery in Alzira, Spain. *J. Build. Perform. Simul.* **2018**, *11*, 369–390. [CrossRef]
22. Selfridge, R.; Cook, J.; McAlpine, K.; Newton, M. Creating Historic Spaces in Virtual Reality using off-the-shelf Audio Plugins. In Proceedings of the 2019 AES International Conference on Immersive and Interactive Audio, York, UK, 27–29 March 2019; Audio Engineering Society: York, UK, 2019.
23. Postma, B.N.J.; Katz, B.F.G. Perceptive and objective evaluation of calibrated room acoustic simulation auralizations. *J. Acoust. Soc. Am.* **2016**, *140*, 4326–4337. [CrossRef]
24. Katz, B.F.G.; Poirier-quinot, D.; Postma, B.N.J. Virtual reconstructions of the Théâtre de l ' Athénée for archeoacoustic study. In Proceedings of the 23rd International Congress on Acoustics integrating 4th EAA Euroregio 2019, Aachen, Germany, 9–13 September 2019; Deutsche Gesellschaft für Akustik: Aachen, Germany, 2019; pp. 303–310.
25. Kleiner, M.; Dalenbäck, B.I.; Svensson, P. Auralization—An Overview. *J. Audio Eng. Soc.* **1993**, *41*, 861–875.
26. Mehra, R.; Rungta, A.; Golas, A.; Lin, M.; Manocha, D. WAVE: Interactive Wave-based Sound Propagation for Virtual Environments. *IEEE Trans. Vis. Comput. Graph.* **2015**, *21*, 434–442. [CrossRef] [PubMed]
27. Pelzer, S.; Aspöck, L.; Schröder, D.; Vorländer, M. Interactive Real-Time Simulation and Auralization for Modifiable Rooms. *Build. Acoust.* **2014**, *21*. [CrossRef]
28. Poirier-Quinot, D.; Katz, B.; Noisternig, M. EVERTims: Open source framework for real-time auralization in VR. In Proceedings of the AM '17, London, UK, 23–26 August 2017; ACM: New York, NY, USA; London, UK, 2017.
29. Vorländer, M. *Auralization. Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*, 1st ed.; Springer: Berlin, Germany, 2008; pp. 1–335. [CrossRef]
30. Anderson, J.S.; Bratos-Anderson, M. Acoustic coupling effects in St Paul's Cathedral, London. *J. Sound Vib.* **2000**, *236*, 209–225. [CrossRef]
31. Álvarez-Morales, L.; Lopez, M.; Alvarez-Corbacho, A.; Bustamante, P. Mapping the acoustics of Ripon cathedral. In Proceedings of the 23rd International Congress on Acoustics, Aachen, Germany, 9–13 September 2019; Deutsche Gesellschaft für Akustik: Aachen, Germany, 2019.
32. Boren, B.; Longair, M.; Orlowski, R. Acoustic Simulation of Renaissance Venetian Churches. *Acoust. Pract.* **2013**, *1*, 17–28. [CrossRef]
33. Alonso, A.; Suárez, R.; Sendra, J. The Acoustics of the Choir in Spanish Cathedrals. *Acoustics* **2018**, *1*, 35–46. [CrossRef]
34. Martellotta, F. Subjective study of preferred listening conditions in Italian Catholic churches. *J. Sound Vib.* **2008**, *317*, 378–399. [CrossRef]
35. Martellotta, F. Identifying acoustical coupling by measurements and prediction-models for St. Peter's Basilica in Rome. *J. Acoust. Soc. Am.* **2009**, *126*, 1175–1186. [CrossRef] [PubMed]
36. Andersson, I.; Ljungstedt, S.; Malm, G. *Vadstena Klosterkyrka. 1, Kyrkobyggnaden*; Riksantikvarieämbetet och Kungl. Vitterhets-, Historie-, och Antikvitetsakademien: Borås, Sweden, 1991.
37. Strinnholm Lagergren, K. The invitatory antiphons in Cantus sororum: a unique repertoire in a world of standard chant. *Plainsong Mediev. Music.* **2018**, *27*, 121–142. [CrossRef]
38. Berthelson, B. *Studier i Birgittinerordens Byggnadsskick 1, Anläggningsplanen och dess Tillämpning*; Kungl. Vitterhets-, historie- och antikvitetsakademien: Stockholm, Sweden, 1947.
39. Andersson, A. *Vadstena Klosterkyrka II. Inredning*; Riksantikvarieämbetet och Kungl Vitterhets-, historie- och antikvitetsakademien: Borås, Sweden, 1983.
40. Sigurdson, J.M.; Zachrisson, S. *Aplagårdar och Klosterliljor: 800 år kring Vadstena Klosters Historia*; Artos: Skellefteå, Sweden, 2010.
41. Nyberg, T.S. *Birgittinische Klostergründungen des Mittelalters*; Gleerup: Lund, Sweden, 1965.
42. Bennett, R. *Vadstena klosterkyrka III. Gravminnen*; Riksantikvarieämbetet och Kungl. Vitterhets-, Historie-, och Antikvitetsakademien: Borås, Sweden, 1985.
43. Autio, H.; Bard, D. A statistical method for parameter estimation from Schroeder decay curves. In Proceedings of the INTER-NOISE 2018, The 47th International Congress and Exposition on Noise Control Engineering, Chicago, IL, USA, 26–29 August 2018; INCE-USA: Chicago, IL, USA, 2018.
44. Alonso, A.; Sendra, J.J.; Suárez, R. Sound Space Reconstruction in the Cathedral of Seville for major feasts celebrated around the main chancel. In Proceedings of the Forum Acusticum 2014, Krakow, Poland, 7–12 September 2014; The Polish Acoustical Society: Krakow, Poland, 2014.
45. Sancta Birgitta. *Reuelaciones Extrauagantes*; Almqvist & Wiksell: Stockholm, Sweden, 1956.
46. Klemming, G.E. (Ed.); Lucidarium. In *Heliga Birgittas Uppenbarelser*; P. A. Nordstedt & Söner: Stockholm, Sweden, 1884; Volume 5.
47. Ugglas, C.R.A. En svensk korsväg med målade stationsbilder, del 1. *Fornvännen* **1938**, *33*, 230–272.
48. Sancta Birgitta. *Opera Minora, Vol 1: Regula Saluatorie*; Almqvist & Wiksell: Stockholm, Sweden, 1975.

49. Boren, B.; Abraham, D.; Naressi, R.; Grzyb, E.; Lane, B.; Merceruio, D. Acoustic Simulation of Bach's Performing Forces in the Thomaskirche. In Proceedings of the 1st EAA Spatial Audio Signal Processing Symposium, Paris, France, 6–7 September 2019; Sorbonne Université: Paris, France, 2019.
50. ODEON Room Acoustics Software. 2020. Available online: https://www.geonoise.com/odeon-room-acoustics-software/ (accessed on 26 December 2020).
51. Postma, B.N.J.; Katz, B.F.G. Creation and calibration method of acoustical models for historic virtual reality auralizations. *Virtual Real.* **2015**, *19*, 161–180. [CrossRef]
52. ISO 3382-1:2009. *Acoustics—Measurement of Room Acoustic Parameters—Part 1: Performance Spaces*; Technical Report; ISO: Geneva, Switzerland, 2009.
53. Pilch, A. Optimization-based method for the calibration of geometrical acoustic models. *Appl. Acoust.* **2020**, *170*. [CrossRef]
54. Vorländer, M. Computer simulations in room acoustics: Concepts and uncertainties. *J. Acoust. Soc. Am.* **2013**, *133*, 1203–1213. [CrossRef]
55. Suárez, R.; Alonso, A.; Sendra, J.J. Archaeoacoustics of intangible cultural heritage: The sound of the Maior Ecclesia of Cluny. *J. Cult. Herit.* **2016**, *19*, 567–572. [CrossRef]
56. Alonso, A.; Suárez, R.; Sendra, J.J. Virtual reconstruction of indoor acoustics in cathedrals: The case of the Cathedral of Granada. *Build. Simul.* **2017**, *10*, 431–446. [CrossRef]

Paper B

buildings

IMPACT FACTOR 3.8

CITESCORE 3.1

Article

The Influence of Different Scattering Algorithms on Room Acoustic Simulations in Rectangular Rooms

Hanna Autio, Nikolaos-Georgios Vardaxis and Delphine Bard Hagberg

Special Issue

Computational and Experimental Evaluation of Architectural Acoustics in Enclosures

Edited by

Assoc. Prof. Dr. Cheol-Ho Jeong

MDPI

https://doi.org/10.3390/buildings11090414 buildings

MDPI

Article

The Influence of Different Scattering Algorithms on Room Acoustic Simulations in Rectangular Rooms

Hanna Autio *, Nikolaos-Georgios Vardaxis * and Delphine Bard Hagberg *

Division of Engineering Acoustics, Lund University, Box 118, 221 00 Lund, Sweden
* Correspondence: hanna.autio@construction.lth.se (H.A.);
nikolaos-georgios.vardaxis@construction.lth.se (N.-G.V.); delphine.bard@construction.lth.se (D.B.H.)

Abstract: Raytracing is a widespread tool for room acoustic simulations, and one of its main advantages is the inclusion of surface scattering. Although surface scattering has been acknowledged as a central aspect of accurate raytracing simulations for many years, there is ongoing research into its effects and how to implement it better. This study evaluates three different algorithms for surface scattering in raytracers, referred to as on–off scattering, perturbation scattering, and diffuse field scattering. Their theoretical foundation is discussed, and the physical accuracy of the resulting simulations is evaluated by comparing simulated room acoustic parameters to measurements. It is found that the choice of surface scattering algorithm has a significant impact on the simulation outcomes, both in terms of physical accuracy and in terms of usability. Additionally, there are differences in the parametrization of surface scattering depending on the algorithm chosen. Of the three tested algorithms, the most commonly used algorithm (on–off scattering) seems to have the best properties for simulations.

Keywords: acoustic raytracing; surface scattering; room acoustic simulation

Citation: Autio, H.; Vardaxis, N.-G.; Hagberg, D.B. The Influence of Different Scattering Algorithms on Room Acoustic Simulations in Rectangular Rooms. *Buildings* **2021**, *11*, 414. https://doi.org/10.3390/buildings11090414

Academic Editor: Cheol-Ho Jeong

Received: 31 July 2021
Accepted: 9 September 2021
Published: 17 September 2021

1. Introduction

Raytracing and other geometrical acoustics (GA) techniques have been some of the most popular sound simulation tools for decades [1,2], after first making their appearance in the 1960s [3]. In the early 1990s, the necessity to account for surface scattering in such models became an established fact [4–6]. Although the importance of surface scattering was acknowledged already almost 30 years ago, there is no firm consensus on what the best algorithm for scattering in raytracers is. Software based on GA techniques exist in several iterations, such as ODEON [7], CATT acoustics [8], RAMSETE [9] or RAVEN [10], and there are significant differences in how they implement scattering. In addition, several alternatives are typically presented in overviews [2,11,12]. Furthermore, there is still research into what effects the scattering coefficient has on room acoustic parameters using different scattering algorithms [13,14]. This study aims to shed new light on the differences between various scattering algorithms by isolating their effects and evaluating their functionality from new perspectives.

Raytracing as a concept defines both an underlying model for the acoustic field and the simulation tool used to estimate the predictions made by the model. A brief introduction is made here, and more detailed descriptions can be found in many textbooks [1,2,12]. The basic modeling assumption in raytracers is that the acoustic energy emitted from a sound source can be accurately modeled using an infinite set of acoustic particles, or rays, which follow plane wave propagation laws. The rays carry acoustic energy which is decreased by surface absorption or attenuation in the medium, and they typically carry no phase information. By the plane wave model, they are reflected by surfaces in the specular (mirror-like) direction. This model is approximately valid for frequency ranges where the wavelength is significantly smaller than the dimensions of the enclosure, and modal effects are not dominant. The Schröder frequency [15] can be used as a lower limit for when these assumptions might be valid. The acoustic response predicted by this model is calculated by a Monte Carlo simulation, where a random, finite subset of acoustic ray paths are sampled and used as an estimate of the full response. If enough rays are used in the calculation, the calculated response will not vary significantly due to random fluctuations and be very close to the response predicted by the underlying model.

Within this framework, surface scattering can be defined using the suggestion by Morse and Ingard [16]. By this definition, surface scattering is the discrepancy between the ideal reflection of a plane wave from a plane, rigid surface, and the real reflected wave. Due to the modeling assumptions in raytracers, the ideal plane wave reflection is easily defined as the specular reflection, adjusted for the energy absorption by the surface. Models for surface scattering should then aim to compensate for the discrepancy between this model and a more accurate model for the distribution of reflected energy.

As there are variations in the terminology used to discuss surface scattering, a few explanations and motivations for the terms used in this study are provided. The terms *diffuse reflection*, *diffusion*, and *scattering reflection* are all used in the context of surface scattering and raytracing, sometimes more or less interchangeably. In this study, the term *scattering reflection* is preferred and refers to any reflection that deviates from the ideal plane wave prediction (in the raytracer case, the specular reflection). The term *diffusion* refers to processes which lead to a more diffuse sound field, i.e., a more uniform distribution of acoustic energy [1]. The term *diffuse reflection* is not used in the sequel.

There are several effects that may cause reflected acoustic energy to deviate from the specular prediction. In raytracing, the plane wave, energetic model is itself an approximation of sound propagation from a point source and may be inaccurate. In general, if the acoustic ray has travelled a distance r, such that $kr \gg 1$, where k is the wavenumber, the approximation is acceptable [1]. Other sources of scattering may be impedance variations across a surface [1]. Surface elements may be excited and re-radiate acoustic energy [16] in a way that is not accounted for in the rigid surface model. Finally, rough surfaces may introduce scattering by spatial dispersion, temporal dispersion due to unequal path lengths, or edge effects such as diffraction [1,2,12]. All these phenomena may cause the real distribution of reflected energy to deviate from the specularly reflected ray model in the basic raytracer.

Raytracers generally approach scattering reflections by formulating an alternative model for a ray's energy distribution after reflection. A scattering algorithm is developed based on this model, and the algorithm's physical accuracy is thus determined by the accuracy of the underlying model. However, selecting a model depends on more factors than physical accuracy. It must also be possible to adapt the mathematical model to raytracers, either by allowing for straight-forward calculations in a different framework or by being appropriate for continued raytracing. Although there are some examples where the first of these options have been implemented [4,5], the second option is more frequently described in the literature [1,2,11,12]. Continued raytracing requires that the model for the reflected energy can be discretized into acoustic particles or rays, which follow plane wave propagation laws. Such models may be constructed based on solutions to the wave equation for that particular surface (which can take all the phenomena discussed above into account); by using a more simplified but specialized model; or by using a general model, such as a Lambertian distribution. When a raytracing-ready model for the energy distribution has been established, the raytracing simulation may continue by sampling the distribution in the same way that the acoustic energy emitted from the sound source is sampled. In addition to technical limitations, the choice of model should lead to a scattering algorithm that is easy and efficient to use. In all, this has led to the Lambertian distribution being the most popular basis for mathematical models of the energy distribution after scattered reflections [2,11,12].

There are two standard coefficients that aim to quantify surface scattering from real surfaces. These are the *directional diffusion coefficient* defined in ISO 17497-2 [17] and the *random-incidence scattering coefficient* defined in ISO 17497-1 [18]. The directional diffusion coefficient measures the similarity between the distribution of the reflected energy and a Lambertian distribution. The random-incidence scattering coefficient measures how much of the incident energy is reflected in a direction other than the specular direction when the incident energy is randomly distributed in space. A more detailed description of the two coefficients can be found in [19]. Although these two coefficients describe aspects of the energy distribution after reflection, they are not by themselves enough to formulate a complete scattering algorithm for raytracers. They may however be useful as tools to evaluate existing models or as input values for parametric models. In raytracers in general, it is often difficult but important to find the appropriate material parameters [20–22] and the relevance of standard coefficients as input parameters should be considered in the choice of scattering algorithm.

The variations in terminology, models, and standards show a need for more research into surface scattering and how it should be implemented in raytracers. Several papers in recent years examine the effects on simulated room acoustic parameters as the scattering parameter varies and has found that changing the scattering parameter has different effects in different spaces [13] and for different algorithms [14]. A deeper understanding of these results can be achieved by further discussing the physical interpretations of increasing or decreasing the scattering coefficient, and by isolating the effects of different scattering algorithms. These are both goals of the study presented in this paper.

This study describes and evaluates three different raytracing scattering algorithms based on their physical accuracy, simulation speed, and *usability*. Raytracers are popular partly because they are fast and easy to use, and it is thus important that the scattering algorithms introduced do not violate any of those properties. In order to evaluate the usability of the algorithms, their predictability and sensitivity to parameter errors are reviewed. This study also implements the three scattering algorithms within the framework of one raytracer, isolating the influence of the scattering algorithm itself and separating this study from previous research [5,14].

In this paper, the three scattering algorithms are firstly presented. Then, the methods used for comparing them are explained more in detail, and finally the results of the evaluation are presented and discussed.

2. The Algorithms Implemented in This Study

In this section, the three scattering algorithms subject to the study are described. The three algorithms are herein referred to as *on–off scattering*, *perturbation scattering*, and the *diffuse field algorithm*. They are based on three different models for the distribution of energy after a scattering reflection. The on–off scattering and perturbation scattering algorithms are implemented entirely within the raytracing framework, whereas the diffuse field algorithm uses external calculations for parts of the simulation. The reasons for these discrepancies and their implications are discussed. A schematic image of the properties of on–off and perturbation scattering is shown in Figure 1.

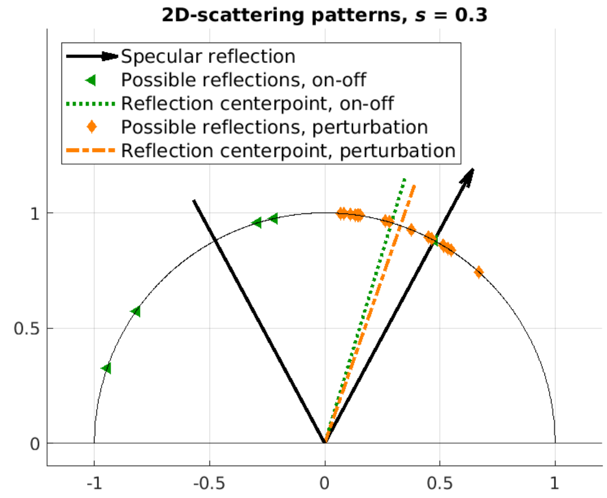


Figure 1. 2D view of some possible reflected rays in the on–off and perturbation scattering algorithms. Center points for the aggregate reflected energy are shown for both algorithms. Approximately 70% of the rays reflected using the on-off algorithm will be reflected in the specular direction.

2.1. On–Off Scattering

The algorithm herein referred to as on–off scattering is possibly the most common way of implementing surface scattering in raytracers, and it is described in several sources [1,11,12]. It is related to the random-incidence scattering coefficient defined in ISO 17497-1 [18] and uses a Lambertian distribution for the scattered energy.

In the underlying mathematical model, the acoustic energy reflected from a surface consists of a combination of energy reflected specularly and energy scattered according to a Lambertian distribution. The ratio of energy that is scattered is described by a value $s_o \in [0,1]$, called the on–off scattering parameter. All remaining energy is reflected in the specular direction. The value of s_o needs to be defined for each surface separately.

This model is easy to adapt to an algorithm suitable for raytracing simulations. For each ray that hits a surface, it is reflected either in the specular direction or in a random direction generated from a Lambertian distribution. The choice of whether a ray should be reflected specularly or randomly is made so that approximately s_o of all rays intersecting this surface are scattered, which ensures that also a ratio s_o of the total energy incident on this surface is reflected according to the Lambertian distribution. An example of the pattern of reflection directions for $s_o = 0.3$ is shown in Figure 1. The name used for the algorithm comes from the on–off behavior for each ray.

The parameter s_o easily relates to the random-incidence scattering coefficient, as defined in the standards [18,19]. The random-incidence scattering coefficient s_{ISO} measures the ratio of the total reflected energy that is not reflected in the specular direction. A ratio of $1 - s_{\text{ISO}}$ is consequently reflected in the specular direction. In the on–off scattering algorithm, $1 - s_o$ of the acoustic energy is reflected in the specular direction and in this way s_o and s_{ISO} are similar. However, there are still differences in the underlying model. When measuring the random-incidence scattering coefficients, no information is obtained regarding the distribution of the energy which is scattered. In contrast, the on–off scattering algorithm uses a Lambertian model for this energy. The Lambertian distribution is a reasonable best guess but cannot be assumed to be generally valid, and some care should be taken before using s_{ISO} as the on–off scattering parameter.

2.2. Perturbation Scattering

Similarly to on–off scattering, perturbation scattering is implemented entirely within the raytracing framework and parametrized by a single value. It is commonly presented as an alternative to the on–off scattering algorithm [11,12] but does not seem to be as frequently used.

In the mathematical model underlying perturbation scattering, acoustic energy after a scattering reflection forms a cone around the specular direction. The cone's width is determined by the perturbation scattering parameter, $s_p \in [0, 1]$. The larger the parameter, the wider the cone. For $s_p = 1$ the distribution defined by the perturbation scattering algorithm will coincide with the Lambertian distribution.

Similarly to on–off scattering, this model is very suitable for raytracing simulations. As a ray intersects a surface, it is reflected in a direction within the cone. The direction is determined by applying a random perturbation (hence the name perturbation scattering) to the direction of specular reflection. The perturbation is generated from a Lambertian distribution and scaled by the perturbation scattering parameter s_p. The direction of the reflected ray d_r can thus be described as $d_r = (1 - s_p)d_s + s_p d_p$, where d_s is the specular direction and d_p is the random perturbation. An example of the pattern of reflections for $s_p = 0.3$ is shown in Figure 1.

The perturbation scattering parameter s_p is distinct from the random-incidence scattering coefficient s_{ISO}. The standard random-incidence scattering coefficient defines how much of the reflected energy is directed away from the specular direction, whereas the perturbation scattering parameter regulates how much the reflected energy in general deviates from the specular direction. This random-incidence scattering coefficient can consequently not be assumed to be appropriate for use as perturbation scattering parameter.

2.3. *The Diffuse Field Algorithm*

The diffuse field algorithm differs from both on–off scattering and perturbation scattering in that it does not rely entirely on raytracing. Although these types of algorithms are less common, there are some previous implementations [4,23]. Similarly to the on–off and perturbation scattering algorithms, the diffuse field algorithm is parametrized using a single value $s_d \in [0, 1]$, the diffuse field scattering parameter.

The mathematical model for the diffuse field algorithm relies on the concept of a diffuse acoustic field and interprets surface scattering as an aspect of diffusion. As a sound wave is reflected from a surface, it is assumed that the reflected energy is partly reflected according to the plane wave assumption and partly converted into "diffuse energy" and added to the diffuse field. Energy in the diffuse field is assumed to be uniformly distributed within the volume [1]. The plane-wave segment of the model is easily implemented within a raytracer, and the propagation paths found in this way coincides with statistical implementations of image source models. In each reflection, a ray's energy is adjusted according to surface absorption and the diffuse field scattering parameter, so that $e_r = e_i(1 - \alpha)(1 - s_d)$, where e_r is the energy carried by the reflected ray, e_i is the energy of the incident ray, α is the surface's absorption parameter and s_d the diffuse field scattering parameter. The energy added to the diffuse field, e_d is determined by $e_d = e_i(1 - \alpha)s_d$.

The energy transferred to the diffuse field is recorded separately and assumed to be evenly distributed in the entire volume quickly. After that, it decays exponentially. According to theoretical models for diffuse acoustic fields, the energy from the diffuse field that is received at the detector E_d can be calculated as

$$E_d = A_d \frac{c}{4} w, \tag{1}$$

where A_d is the total surface area of the detector, c is the speed of sound and w is the energy density [1]. The energy density is estimated as $w = \frac{E_{\text{diff}}}{V}$, where E_{diff} is the total energy in the diffuse field and V is the volume of the simulated space. The energy contained within the diffuse field varies over time, as it decays exponentially and more energy is added from surface scattering.

A complication with this model is that energy may be detected in a non-physically short amount of time. As soon as energy is transferred to the diffuse field, no matter where the reflection causing it occurs, it can be detected anywhere in the simulated space. To counteract this, a method suggested by Lam is implemented [23]. Energy is added to the diffuse field after a delay corresponding to the time needed to travel the mean free path.

The decay rate of the diffuse field may be determined in several ways, for example by using theoretical models such as Sabine's or Eyring's formulae. In this case, it is estimated using the raytracer. Several rays are emitted into the modeled space and are reflected in a random direction at each surface. After a certain number of reflections, they are assumed to be well-mixed and approximately uniformly distributed, thus emulating the diffuse field. This step is called a burn-in. After this time, the rays start losing energy according to the absorption parameters of the modeled space and air absorption. The time until the energy of each ray decreases below some pre-defined threshold is recorded and used to estimate the decay rate of the diffuse field. Since the decay rate of the diffuse field is constant for the entire acoustic volume, this step is only needed once for any number of simulations in a given space. In addition, the total number of rays needed is relatively small for two reasons. Firstly, only one parameter needs to be estimated. Secondly, the result for each ray is recorded separately, which allows for statistical outlier detection that stabilizes the estimate. In this study, the decay rate is estimated using 10% of the total number of rays, and the remaining rays are used for the traditional raytracing simulation.

After the raytracing simulation, the results from simulations using the diffuse field algorithm thus consist of three data sets. (1) A reflectogram showing the detected energy from the traditional raytracing simulation; (2) Decay data obtained from the diffuse field decay rate estimation step; and (3) Data showing how much energy is transferred to the diffuse field in each time segment. The energy contained in the diffuse field is determined for each time segment using datasets 2 and 3. Of this energy, the ratio determined by Equation (1) is added to the reflectogram to obtain a full energetic impulse response. This step is performed externally and considered part of the post processing.

In the diffuse field algorithm, surface scattering is parametrized using a single value: The diffuse field scattering parameter s_d. It can be interpreted as the amount of diffusion caused by reflection in a given surface. However, since the concept of a diffuse field itself is an idealization of real sound fields, it is unclear whether this is a physical property that can be measured. The diffuse field scattering parameter is not related to either the random-incidence scattering coefficient or the directional diffusion coefficient (defined in ISO 17497-1 and 2, respectively [17,18]).

The diffuse field algorithm and its underlying model can be interpreted as a combination of an image source model and a diffuse field model. With this interpretation, the diffuse field scattering parameter acts akin to a weighting coefficient, determining the relative influences of the diffuse field and image source models in the simulation.

2.4. Differences and Similarities between the Algorithms

The on–off scattering algorithm and the perturbation scattering algorithm are very similar in many aspects. From a practical standpoint they only differ in the directions in which rays are reflected after hitting a surface. These types of algorithms can easily be constructed by finding new distributions for the reflected rays, possibly based on more complex or physically-based models of various surfaces. This is a frequently suggested modification, which is theorized to improve the accuracy of the simulation as a whole [11,12]. A full comparison of on–off and perturbation scattering may shed light on the impact of the model for the reflected energy, and show whether better models can be expected to lead to significant improvement.

The diffuse field algorithm, on the other hand, is quite dissimilar to the previous two algorithms. The theoretical model it is based on is different in itself, and it relies partly on numerical calculations for parts of the simulations. The numerical calculations may mean that the algorithm is less affected by random noise and thus more reliable. Furthermore, if the calculations are sufficiently optimized, it may be faster to use than running a full raytracing simulation even considering the time needed to determine the diffuse field decay rate. Comparing the diffuse field scattering algorithm to the on-off and perturbation scattering algorithms is expected to help show the differences between hybrid methods and fully raytracing-based algorithms.

All three algorithms described are parametrized using a single scattering parameter, but the physical and numerical significance of this parameter differs between them. Consequently, it cannot be expected that a surface should have the same on-off, perturbation, and diffuse field scattering parameters. Of the three, only the on–off scattering parameter is easily related to one of the standardized coefficients, namely the random-incidence scattering coefficient [18]. For the diffuse field algorithm, in particular, the relation between its scattering parameter and physical properties of surface scattering is hard to define. Instead, it acts as something of a weighting factor between two models for sound propagation, the energetic plane wave model used in raytracers and the purely diffuse field model.

3. Evaluating the Scattering Algorithms

As mentioned in the introduction, the scattering algorithms examined in this study are evaluated in three different aspects:

- Physical accuracy;
- Simulation speed;
- Usability.

In this section, the methodology used for comparing algorithms is presented and discussed.

3.1. Physical Accuracy

The physical accuracy of the scattering algorithms was evaluated by comparison between simulations and measurements. Measurements were carried out in the lower chamber of the impact sound lab in the Division of Engineering Acoustics, LTH (Lund University), shown in Figure 2. It has a total volume of approximately 95 m^3 and was measured in two conditions. The first condition, herein referred to as case A, is a highly reverberant condition where all surfaces are acoustically hard. The second condition, case B, is similar but has introduced a patch of highly absorptive material in the ceiling. It is expected that the concentrated absorption in case B should correspond to a higher sensitivity to the scattering algorithm and the scattering parameter, whereas case A acts as a useful baseline. Further details on the measurements are presented in Section 4.

Simulations and measurements were compared using three standard room acoustic parameters, reverberation time (T_{20}), early decay time (EDT), and speech clarity (C_{50}), as defined in [24]. The reverberation time was included as it is a ubiquitous measure often used as a design parameter, and it is thus crucial that simulation software can predict it accurately. The EDT has been shown to be closely related to the perceived reverberation, and it is consequently important that it can be simulated correctly. Finally, clarity for speech or music (C_{50} and C_{80}, respectively) is a perceptually important aspect of the sound field. Since the measured space is relatively small and, thus, more likely to be used for speech presentation, it was concluded that C_{50} was more relevant. All three room acoustical parameters were compared using the measured just noticeable difference (JND, as defined in [24]), and values that are within 1 JND are assumed to be perceptually similar.

The simulations were performed using absorption and scattering parameters estimated by several methods. It is a common issue in room acoustic simulations that the material parameters of existing surfaces are unknown [20–22] and must be calibrated somehow to achieve a good match between measurements and simulations. Generally, this holds for both scattering parameters and absorption parameters. In this study, it is assumed that the scattering parameter may vary between the algorithms, whereas the absorption parameter corresponds to the sound absorption coefficient as defined in [25] and is the same for all algorithms. Consequently, the scattering parameters were estimated for each of the algorithms individually, as described in Section 4.2, and the absorption parameters were estimated for all algorithms simultaneously using two different methods. The only material for which the absorption coefficient was available was the absorptive material in case B, and table values from the manufacturer were used for this material.

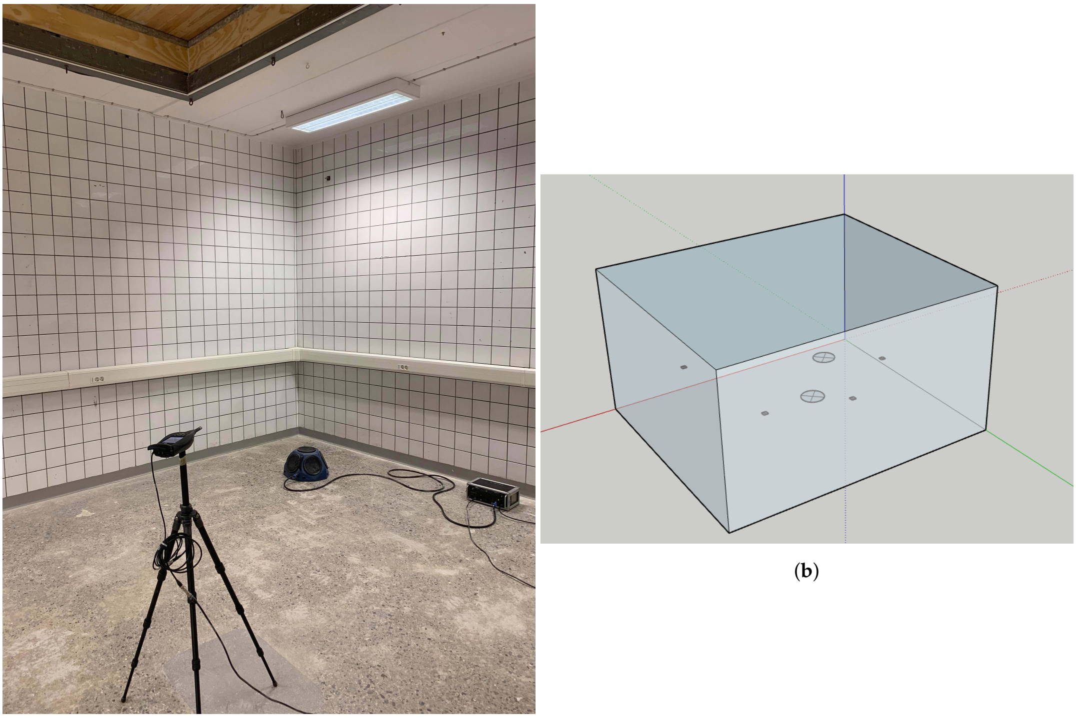

(**a**)
(**b**)

Figure 2. (**a**) Photo and (**b**) model of the space used for measurements and simulation. Two speaker positions and four microphone positions are marked in (**b**).

For middle to high frequencies, the tested algorithms themselves were used to calibrate the absorption parameters. Simulations were run for all algorithms using a wide range of scattering parameters, and spatial average room acoustic parameters were extracted. The results from each algorithm were compared to measurements to see if there was any one scattering parameter value for which all room acoustic parameters were within 1 JND of the measurements. When such a value existed for all algorithms, it was assumed that the absorption parameters were sufficiently accurate.

For frequencies 500 Hz and below, a different method for estimating the absorption parameters were used. Since raytracers are known to perform poorly for frequencies below the Schröder frequency, they were not considered appropriate tools for estimating a value corresponding to a physical property. Instead, a mode-based method presented by Meissner [26] was used. This method estimates the EDC for low frequencies in rectangular spaces with hard, reflective walls. Modal damping factors calculated from the specific wall conductances are used to model the energy decay over time by summating over eigenfrequencies. The EDC from Meissner's calculations was compared to the measured EDC, and the reverberation time was extracted and compared to the measured value. The specific wall conductances were tuned until the calculations were close to the measurements. The absorption coefficients were estimated from the specific wall conductances [15,26] and used as absorption parameters in the simulation.

The algorithms were evaluated both by comparison of spatially averaged room acoustic parameters and based on their ability to predict spatial variations of the sound field. Spatial variations were examined by studying measured room acoustic parameters for each of the positions shown in Figure 2. These were compared to simulated values using the same positions. Ideally, any variations between positions should be similar for both sets of results.

3.2. Simulation Speed

The simulation time for the three algorithms was examined to identify any significant differences between them. In the comparison, only the time used for raytracing has been included, although the diffuse field algorithm requires significantly more external calculations and post processing, increasing the total time consumption. This choice was made since this study focuses on raytracing algorithms in particular.

The time consumption of raytracers generally depends on the number of rays used in the simulation. This, in turn, is governed by the desired level of precision and accuracy. If the number of rays is sufficiently large, the simulation results are consistently very close to the predictions made by the underlying model, whereas too few rays lead to results that may vary significantly between calculations or deviate significantly from the predictions made by the model. How many rays are needed depends on the modeled space [2] including the material parameters and on the algorithm used. In this study, the number of rays needed in each of the three algorithms was evaluated separately. In this way, the time consumption for each algorithm could be measured using the appropriate number of rays.

The sufficient number of rays was determined by reviewing how much the simulation results deviated from the prediction made by the underlying model. This was evaluated by defining the maximum simulation error, e_{Sim}. The e_{Sim} is the largest deviation between a baseline value and a simulated value for a given number of rays. The baseline value is defined as the average result of a set of simulations with many rays, and is thus close to the value predicted by the underlying model. The e_{Sim} was defined for each of the room acoustic parameters mentioned in Section 3.1, according to Equation (2).

$$e_{\mathrm{Sim}} = \max_{p \in P,\, l \leq L} |x_{p,l} - \bar{x}_p|, \tag{2}$$

where $x_{p,l}$ is the simulated value of a room acoustic parameter for source and receiver position p and simulation number l. $\bar{x}_p$ is the baseline value of the selected room acoustic parameter for that position, and P is the set of all positions. L is the total number of simulations run with that specific number of rays. Note that e_{Sim} only measures the reliability of the simulation. A small e_{Sim} indicates that a simulation using this number of rays accurately finds the value predicted by the underlying physical model, not whether that model is accurate or not.

When the e_{Sim} fell below 1 JND for all three room acoustic parameters and octave bands, it was assumed that the number of rays was sufficiently high for this algorithm. The time consumption for a raytracing simulation using this algorithm and this number of rays was then measured and compared to the other algorithms.

3.3. Usability

Finally, the usability of the three algorithms was examined by studying the variations in simulation outcome for varying input scattering parameters. Three aspects were evaluated: whether small variations in input value leads to small changes in simulation outcome; whether changes in the input parameter have a predictable effect; and whether reasonable input values yield reasonable simulation results.

To evaluate how the simulation results vary based on variations in the scattering parameter, a set of simulations using a range of scattering parameter values were performed. According to previous research, changes in the scattering parameter have a larger impact when its value is small [10,13]. Consequently, a finer resolution of scattering parameter values was used for smaller values. By evaluating the room acoustic parameters for each of the scattering parameter values used, an estimate of the influence of the scattering parameter was obtained.

The results from changing the scattering parameter in the three algorithms can be compared to results from previous research to determine whether they are predictable in the sense of behaving similarly to other algorithms. In general, increased scattering parameter lead to either a decreased reverberation time [10,13] or have no effect on it [14,21]. The discrepancies might be due to differences in the spaces being modelled. Research on relatively small, rectangular spaces found that reverberation decreased [10,13]. For the test space in this study, it is thus expected that increases in the scattering parameter should decrease the reverberation time. Regarding clarity, previous research is again inconclusive but indicate that increased scattering has no significant impact [14] or leads to reduced clarity [21]. However, both these results are produced for relatively complex test spaces, and the results in this study may be different.

Reasonable values for the scattering parameter are obtained from the research above, engineering handbooks for some commercial software [27,28] and table values of the random-incidence scattering coefficient [2,12]. For the surfaces in the test spaces, a reasonable initial value of the mid-frequency scattering parameter is about 0.1. Using this scattering parameter, the simulation results should be fairly accurate, although it cannot be expected that 0.1 is the ideal value for all algorithms.

4. Method

In this section, the technical details of the research undertaken in this study is presented.

4.1. Measurements

Impulse response measurements were carried out in the test space using the open-source REW (https://www.roomeqwizard.com/, accessed on 29 August 2021) measurement software installed on a laptop connected to an Audio 8 DJ soundcard from Native Instruments (https://www.native-instruments.com/en/, accessed on 29 August 2021). A Bruel&Kjaer amplifier type 2734 (https://www.bksv.com/en/transducers/acoustic/sound-sources/power-amplifier-2734, accessed on 29 August 2021) connected the soundcard to a dodecahedral loudspeaker, which was used as a sound source. The impulse response was recorded using a Bruel&Kjaer Type 2270 Sound Level Meter and Analyzer (https://www.bksv.com/en/instruments/handheld/sound-level-meters/2270-series/type-2270-s, accessed on 29 August 2021) as a microphone. Two source positions and four microphone positions were used, marked in the digital model in Figure 2. Due to technical limitations and the size of the test space, the sound source was located closer to the floor than described in the measurement standards.

Octave band room acoustic parameters for frequencies in the range 125 Hz to 4000 Hz were extracted using the open-source MATLAB package ITA-acoustics [29]. This includes reverberation time T_{20}, EDT, and C_{50}, as well as the EDC. When applicable, these have been spatially averaged according to the international standard for measurements of these parameters [24].

4.2. Material Parameter Calibration

Meissner's algorithm [26] for estimation of the EDC was implemented in MATLAB. As mentioned in Section 3, this method is based on a summation of the contributions from the room's eigenmodes, subject to damping factors determined by the conductance of the walls. Estimates for case A, the space with hard walls, were obtained for octave bands centered on 125 Hz, 250 Hz, and 500 Hz by summation of all the eigenmodes in the corresponding octave bands. The reverberation time T_{20} was extracted from the EDC.

The Meissner algorithm was used to tune the absorption parameter for the three octave bands in question. This was done iteratively. Initial guesses of the specific wall conductances were set to 0.0053 for all surfaces, corresponding to a random incidence absorption coefficient (and, thus, absorption parameter) of 0.04. Then, this value was adjusted for each of the surfaces until the estimates of the reverberation time were within 1 JND, and good visual correspondence between the measured and the calculated spatial mean EDC was obtained. The absorption parameters corresponding to the found conductances are shown in Table 1.

Table 1. Absorption parameters used in the simulations. The values for the porous absorbers are table values for the random-incidence absorption coefficient, taken from the manufacturer of the product. For bands centered on 125–500 Hz, the absorption parameters are random-incidence absorption coefficients estimated from the conductances used in the Meissner calculations. The remaining values have been estimated by simultaneous tuning for all algorithms, trying to emulate the measurements.

Material	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz
Tiled walls	0.0734	0.056	0.0453	0.04	0.045	0.045
Concrete floor	0.0596	0.0453	0.0453	0.045	0.045	0.045
Ceiling (concrete slab)	0.0665	0.0631	0.0453	0.04	0.04	0.04
Ceiling (porous absorber)	0.25	0.80	0.95	0.95	0.95	0.95

For octave bands centered on 1000 Hz and up, absorption parameters were estimated from initial estimates of $\alpha = 0.04$ for all surfaces except the porous ceiling in case B. This value was tuned iteratively by adjusting the value for all materials individually and running additional simulations for all algorithms. This process continued until the spatial average room acoustic parameters were approximately within 1 JND of the measurements for all algorithms, according to the method described in Section 4.2. Regarding the porous ceiling, absorption parameters for all frequencies (including 125 Hz–500 Hz) were obtained from table values from the manufacturer. The resulting values are shown in Table 1.

In order to find appropriate scattering parameters, a range of scattering values were tried in simulations with each of the algorithms. Scattering parameters were varied over the range $[0.01, 0.99]$, and the tested values are shown in Table 2. From these values, the scattering parameter which yielded the best match between measured and simulated room acoustic parameters were chosen as most appropriate. The best match was evaluated by studying the deviation between simulated and measured values, as calculated by

$$E_{\mathrm{JND}} = \frac{x_{\mathrm{sim}} - x_{\mathrm{meas}}}{1\mathrm{JND}}, \tag{3}$$

where x is one of the spatially averaged room acoustic parameters, and the JND corresponding to that measured parameter is used. Values inside the range $[-1, 1]$ then indicate that the room acoustic parameter is sufficiently well estimated. Using this expression allows for the simultaneous graphical evaluation of all three room acoustic parameters. The scattering parameters were assumed to be constant for all hard surfaces in the test space. Consequently, they were tuned simultaneously for case A and B.

4.3. Simulations

The three scattering algorithms were implemented within *AMRT*, an in-house acoustic raytracer at the Division of Engineering Acoustics at LTH, Lund University. The raytracer is based on the NVIDIA OptiX raytracing engine [30], version 5.1. All models include frequency-dependent air absorption corresponding to a temperature of 20 °C and 70% relative humidity [1]. This is a relatively high humidity for an indoor space, but is reasonable for the test space as it is located in a basement with significant air exchange to the external environment. The outside humidity during the measurements was around 80–90%.

A digital model of the test space in case A (no absorption) and B (localized ceiling absorption) was developed using Trimble SketchUp (www.sketchup.com, accessed on 29 August 2021), shown in Figure 2b. This model was used for simulations.

Table 2. The range of scattering values and number of rays used for simulations. In total, 23 different scattering parameters and 14 different numbers of rays were used.

Scattering Parameters	Number of Rays
0.01	60
0.03	240
0.035	600
0.04	1920
0.045	6000
0.05	18,960
0.055	60,000
0.06	540,000
0.065	600,000
0.07	900,000
0.1	1,200,000
0.13	3,000,000
0.15	4,500,000
0.17	6,000,000
0.2	
0.25	
0.3	
0.35	
0.4	
0.5	
0.7	
0.9	
0.99	

Several simulations using different parameters were run for each algorithm. Here, a single simulation refers to the calculation of the energetic impulse response for the eight combinations of source and listener positions (shown in Figure 2b), in the six octave bands ranging from 125 Hz to 4000 Hz. Simulations were run using 23 different scattering values in the range $[0.01, 0.99]$, shown in Table 2, with 6,000,000 rays. Each simulation was run five times to reduce the effects of random fluctuations. These results were used for calibration of the scattering parameters, as well as evaluation of the physical accuracy and usability of the algorithms.

Additional simulations using the tuned scattering parameters were run to study the time consumption of the algorithms. Five simulations using fourteen different numbers of rays, in the range [60, 6,000,000] (shown in Table 2), were run for each algorithm. These results were used to calculate the maximum simulation error e_{Sim} and measure the time consumption for each of the algorithms.

5. Results and Analysis

In this section, the results of the measurements and simulations are presented and analyzed based on the methodology presented previously. As an initial step, the measurements are presented, and the calibrated absorption and scattering parameters are presented. Subsequently, results pertaining to physical accuracy, simulation speed, and usability are presented.

The measured spatially averaged room acoustic parameters are shown in black in Figure 3. In particular, it should be noted that the reverberation time is lower for the lowest frequencies, especially in case A. This is unexpected, as the materials in the space are generally expected to have lower absorption for lower frequencies, which should lead to a longer reverberation time. Although modal effects could lead to decreased reverberation due to cancellation, such effects are not expected to be dominant in the spatial average results. Instead, it is assumed that there is some transmission of low-frequency acoustic energy through the walls and the concrete ceiling. As these results are used to obtain the absorption parameters used in the simulations (shown in Table 1), the absorption parameters are in general higher for low frequencies than what is typically supplied in engineering tables. Overall, the values of the absorption parameters seem consistent with the very hard, mostly reflective surfaces in case A.

The calibrated scattering values used for simulation are shown in Table 3, obtained using the method described in Section 4.2. For the absorbers in case B, it was found that the scattering parameter had no impact on the simulation outcome. Accordingly, they were assigned a constant value. For the remaining surfaces, the optimal scattering parameter was determined by graphical evaluation of the relative deviation from measurements, as determined by Equation (3). Examples of the graphs used are shown in Figure 4. In many cases, there existed a scattering parameter value, such that all room acoustic parameters were estimated sufficiently well. When this is not the case, scattering parameters for which most room acoustic parameters are well-estimated were selected, if available. However, in some cases, no such values could be found. This occurs for low frequencies and for the perturbation and the diffuse field scattering algorithms, and the relevant values are marked in Table 3. Possible explanations for these issues are described in Section 5.1.

Table 3. Suitable scattering values for the two cases and different algorithms. They have been determined using the method described in Figure 4. Missing values indicate that no scattering value could be found for which the average room acoustic parameters were accurate. Values indicated by † are scattering values that outperform other scattering values in the given case, although some room acoustic parameters are outside 1 JND.

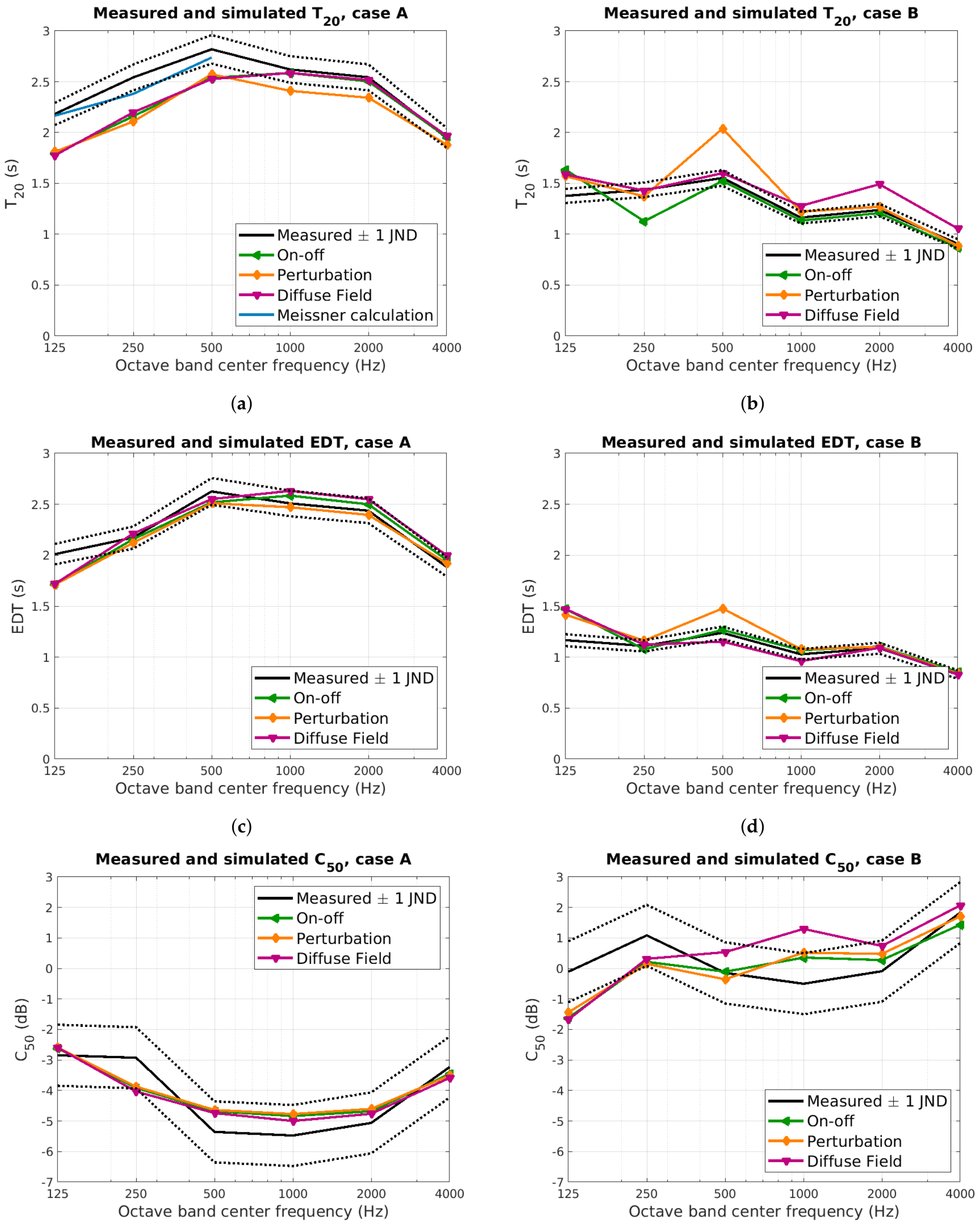
		125 Hz	**250 Hz**	**500 Hz**	**1000 Hz**	**2000 Hz**	**4000 Hz**
Hard surfaces	On-off	-	-	0.05 †	0.13	0.1	0.25
	Perturbation	-	0.09 †	-	0.135 †	0.12 †	0.2
	Diffuse Field	-	-	0.03 †	0.05 †	0.035 †	0.05 †
Absorbers	On-off	-	-	0.1	0.1	0.1	0.1
	Perturbation	-	-	0.1	0.1	0.1	0.1
	Diffuse Field	-	-	0.1	0.1	0.1	0.1

† signifies values that do not fulfill the criteria but outperform other values.

As previously discussed, the scattering parameters have no clearly defined physical counterpart, and their "accuracy" is, thus, hard to evaluate. However, some comments can be made regarding how well they correspond to physical measurements of scattering in general and suggested scattering parameters from engineering tables. Firstly, it is noted that the tuned scattering parameters generally increase as the frequency increases for all algorithms except between 1000 Hz and 2000 Hz. Increased scattering with increased frequency is consistent with typical results when scattering is measured [12] and suggested parametrization in commercial software [28,31]. Secondly, the tuned scattering parameters are relatively small, as expected for the smooth surfaces in the test space.

5.1. Physical Accuracy

To evaluate the physical accuracy of the model, the spatially averaged room acoustic parameters are considered, as shown in Figure 3. In most cases, simulation results are within 1 JND for frequencies above 1000 Hz for all room acoustic parameters. However, both the perturbation algorithm and the diffuse field algorithm results deviates from measurements and other simulated values by more than 1 JND in several cases. This coincides with the frequency bands for which the corresponding scattering parameter could not be tuned. The poor results for the diffuse field algorithm in case B is likely explained in part by issues in the underlying modeling assumptions and how they relate to the modeled space. As described in Section 2, the diffuse field algorithm assumes that part of the acoustic field can be accurately modeled using a diffuse field model. When there is a localized patch of absorbent material, as in case B, there is typically a net flow of energy towards the absorbers and the diffuse field model is invalidated. It is reasonable that the scattering algorithm fails when the underlying modeling assumptions are inaccurate.

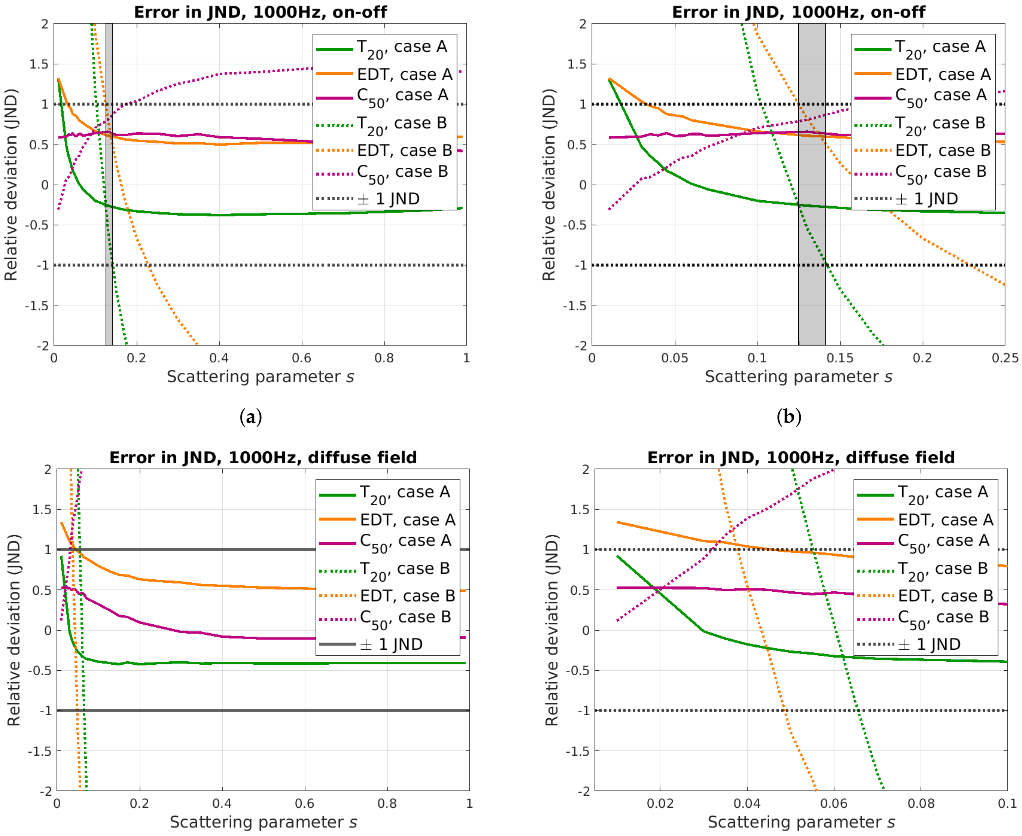


(**a**) (**b**)

(**c**) (**d**)

(**e**) (**f**)

Figure 3. Spatially averaged results from measurements and simulations. The simulated results have been produced using the absorption and scattering parameters shown in Tables 1 and 3. The measured results are presented with dotted lines indicating values within 1 JND. In (**a**), the estimated reverberation time using Meissner's algorithm is shown [26].



(**a**) (**b**)

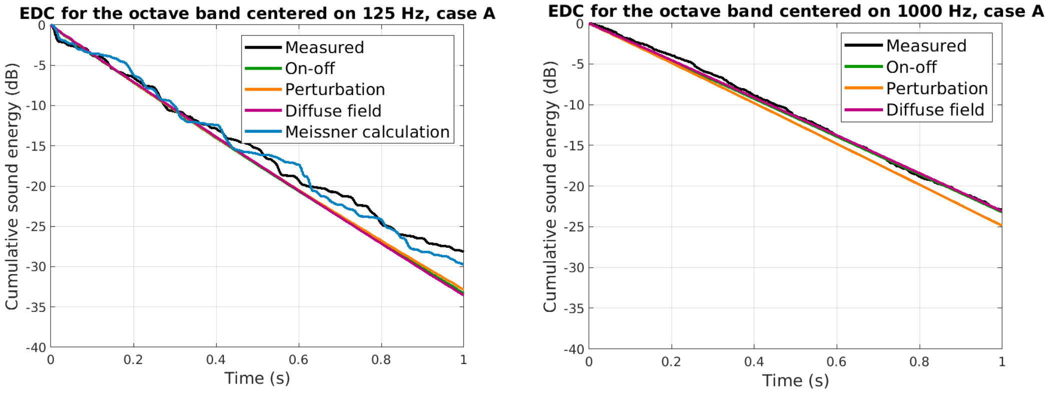
(**c**) (**d**)

Figure 4. Graphs illustrating the procedure of tuning the scattering parameter. The graphs show the deviation between the simulated room acoustic parameters and the measurements at 1000 Hz, normalized to the respective JND, as shown in Equation (3). For the simulations to be sufficiently close to the measurements, all room acoustic parameters should be within 1 JND of the measurements. This occurs within the grey area in Figures (**a**,**b**) showing results for the on–off algorithm, but does not occur simultaneously for all parameters in Figures (**c**,**d**), showing results for the diffuse field algorithm. Note the differences in scale along the x-axis between Figures (**b**,**d**).

The discrepancy between measurements and simulations for octave bands up to 500 Hz can be explained by the issues in modelling modal behavior in raytracers. The fact that the results from the Meissner calculations, shown in Figure 3a, model the sound field more accurately supports this theory. The difference between low and middle to high frequencies is further illustrated by comparing the EDC, as shown in Figure 5.

There are apparent qualitative differences between the measured EDC for 125 Hz and that for 1000 Hz. In the higher frequency band, the measured EDC is almost linear, whereas the measured EDC for 125 Hz is much more dynamic due to lower modal density. The complex behavior in the low frequency case is emulated quite well by the EDC produced from Meissner's calculation, although they are not identical. Regarding the raytracer results, however, there are no such qualitative differences between low and high frequency behavior. The EDC is almost linear in both cases.

The room acoustic parameters so far have been presented as spatial averages, neglecting any spatial variations. The measured and simulated T_{20} and C_{50} for each position are shown in Figure 6. There is none or minimal spatial variation in the measured reverberation time, and this is replicated in the simulations for all algorithms. In case B, it is again evident that the diffuse field algorithm deviates from the measured value and the values predicted from the other algorithms. However, the deviation is very consistent in size, and the spatial variations are thus similar for all three algorithms.



(**a**) (**b**)

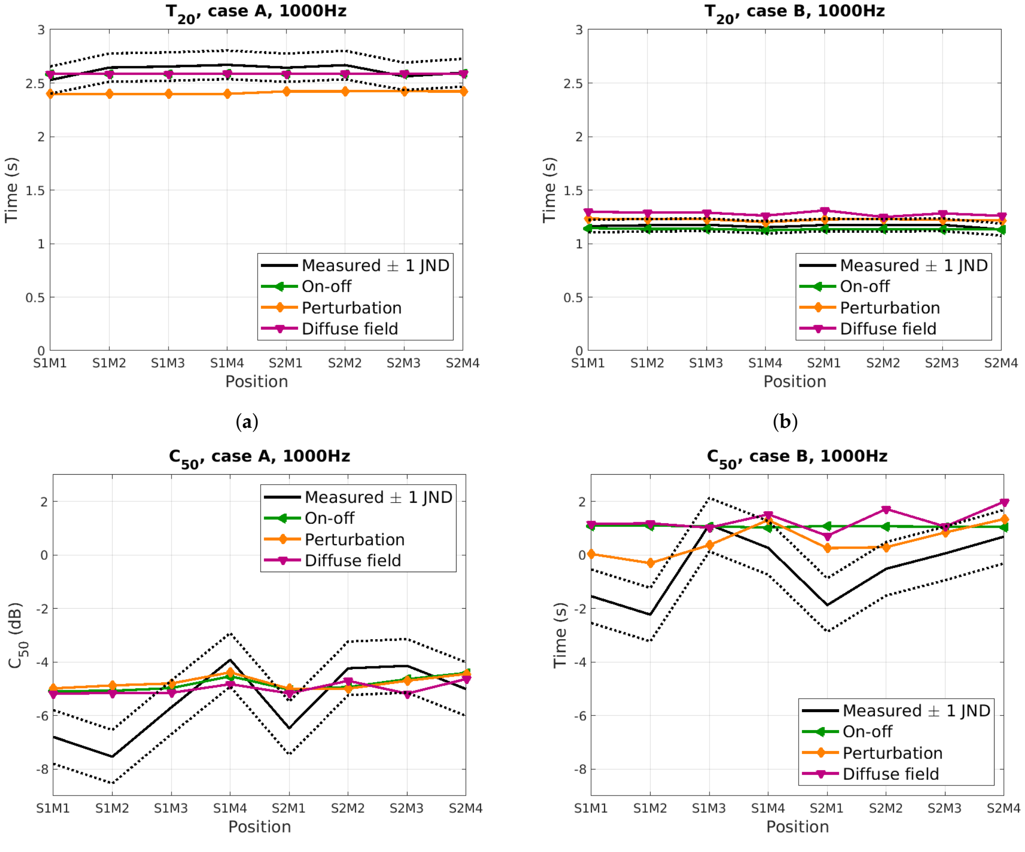
Figure 5. Energy Decay Curves for the octave bands centered on 125 Hz and 1000 Hz in case A, as produced by measurements and simulation, as well as (in the case of 125 Hz) Meissner's algorithm. The results are for source position 1 and listener position 1 in Figure 2. In (**a**), there are significant differences between measurements and Meissner's calculations on one hand and simulations on the other. This may be indicative of influences from room modes. In (**b**), the differences between simulations and measurements are smaller. The measured decay curve in the higher-frequency case is much more linear.

There are more significant spatial variations in the results for C_{50}. This parameter should be more significantly impacted by strong early reflections and thus vary more for different positions. The variations in the measured data are much larger than in the simulated data, although the simulations are mostly within 1 JND of the measurements. There could be several explanations for the discrepancies, including issues both in the measurement process and the simulation process. Measurement noise may lead to random fluctuations in the measured C_{50}. On the other hand, the lack of spatial variations in the simulated C_{50} may reflect shortcomings in the simulations. Since the C_{50} depends on early reflections, it can be heavily influenced by any simplifications made in the modeling process. In fact, as the geometric model for the space is so regular, it is quite expected that the simulated data do not show any spatial variations.

Based on the results shown in Figure 6, none of the scattering algorithms tested seems to outperform the others in emulating spatial variations in the measurements.

5.2. Simulation Speed

As a first step in measuring the time needed for a simulation of each algorithm, the number of rays needed is determined using the maximum simulation error e_{Sim}, as defined in Equation (2) in Section 3.2. The e_{Sim} calculated for T_{20} and C_{50} is shown in Figure 7. Since the results for the EDT are very similar to the results for C_{50}, they are not shown here. As expected, the error decreases as the number of rays increase. The number of rays needed for e_{Sim} to be below 1 JND is presented in Table 4. It should be noted again that e_{Sim} does not relate to physical accuracy, and the results here indicate whether the calculated response matches the underlying model, not whether it matches reality.



(**a**) (**b**)

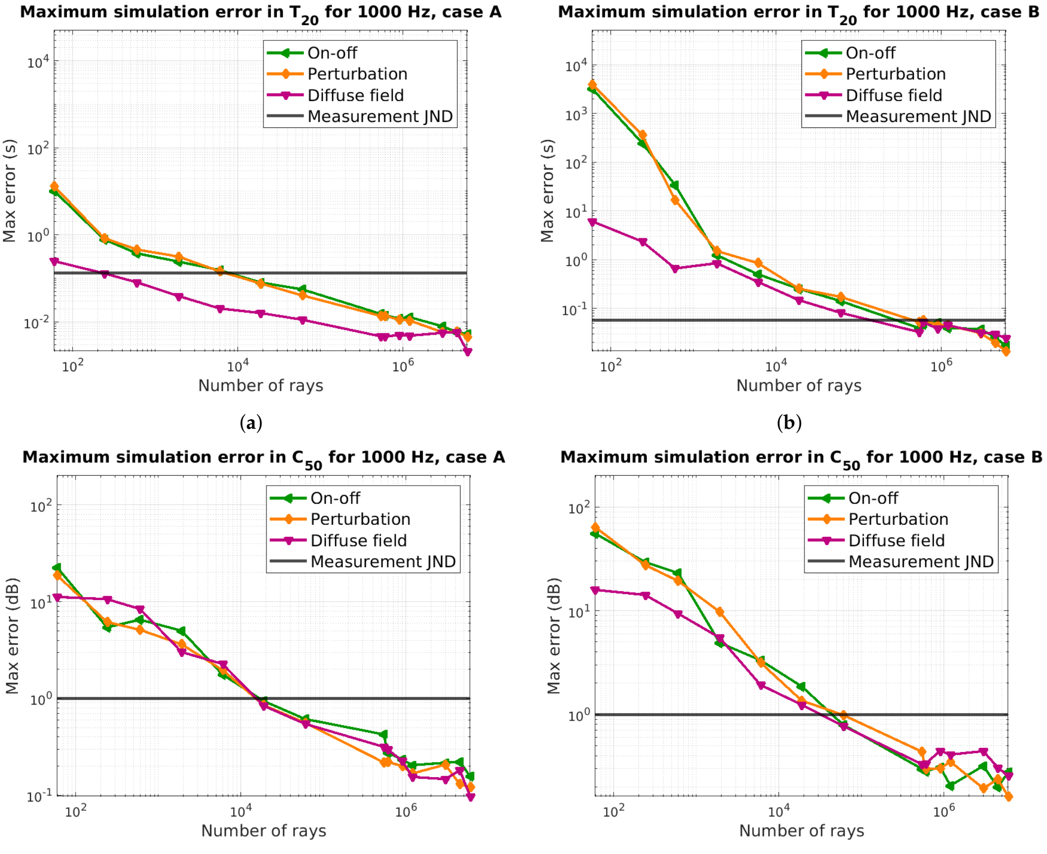
(**c**) (**d**)

Figure 6. Measured and simulated T_{20} (**a**,**b**)and C_{50} (**c**,**d**) for the octave band centered on 1000 Hz, for each source and listener position. There are only small spatial variations for the reverberation time, which is consistent with theory. Similarly to the results seen for spatially averaged measurements, the perturbation algorithm is inaccurate in case A and the diffuse field algorithm in case B. For C_{50}, the measured spatial variations are larger. The variations in the simulated results are minor.

In Figure 7, it is shown that e_{Sim} for T_{20} in case A is smaller for the diffuse field algorithm than the other two. This also occurs for small numbers of rays in case B. The reverberation time is related to the late, more reverberant parts of the sound field, which for the diffuse field algorithm corresponds to the parts of the sound field that are modeled using the diffuse field assumptions. The relatively small simulation error for T_{20} shows that the methods for calculating the decay rate and energy added to the diffuse field are reliable and numerically stable in the chosen implementation.

From the results presented in Table 4, there are no significant differences in the number of rays needed for each of the algorithms. If only reverberation time was considered, it is likely that the diffuse field algorithm would outperform the others, based on the results in Figure 7.

It is also shown in Figure 7 that simulations in case B need more rays before e_{Sim} is small enough. Since theoretical estimates of the standard deviation of raytracing results indicate that the error should decrease as the equivalent absorption area increases [2], these results are somewhat unexpected. Although the standard deviation and the e_{Sim} are not identical, they should generally follow the same tendencies and it is thus expected that e_{Sim} should be smaller in case B. However, the theoretical estimate of the standard deviation requires the absorption to be relatively uniformly distributed in the test space, which is not the case for case B. Instead, the increased error e_{Sim} indicates that the complexity of the sound field has increased as the patch of high-absorption material was added.



(**a**) (**b**)

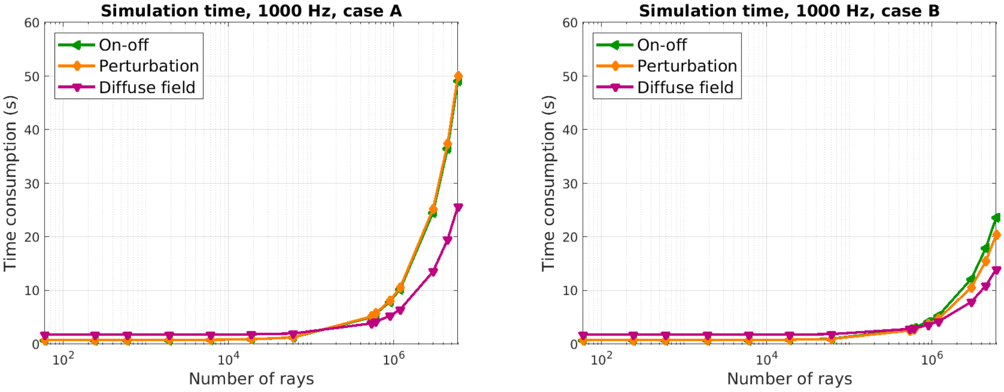
(**c**) (**d**)

Figure 7. The maximum simulation error in T_{20} (**a**,**b**) and C_{50} (**c**,**d**) as it varies for different numbers of rays in the simulation. A limit of 1 JND, as estimated from the measurements, is included in the graph. As expected, the simulation error decreases as the number of rays increase. This decrease is faster for case A.

Table 4. The number of rays needed for the e_{Sim} to be below 1 JND for all the examined room acoustic parameters and the time consumption for such a simulation. More rays are needed for all algorithms in Case B.

	Algorithm	Number of Rays	Time Consumption (s)
Case A	On-off	60,000	1.23
	Perturbation	60,000	1.25
	Diffuse Field	60,000	1.98
Case B	On-off	540,000	2.80
	Perturbation	540,000	2.54
	Diffuse Field	540,000	2.84

The time needed for running a raytracing simulation for a varying number of rays for each of the algorithms is shown in Figure 8. The time measured includes the raytracing for all four listener positions and one of the source positions for all frequency bands. The diffuse field algorithm data includes the raytracing step used to determine the diffuse field decay rate. The results presented in this graph are used together with the results shown in Figure 7 to estimate the time needed for a simulation, shown in Table 4.



(**a**) (**b**)

Figure 8. The average time consumption in seconds for a simulation of case A (**a**) and case B (**b**). The diffuse field algorithm takes longer to run for a small number of rays, but is faster when a large number of rays are used.

Table 4 shows that the on–off and perturbation scattering algorithms have similar time consumption, while the diffuse field scattering algorithms is slower in case A and similar in case B.

Some more comments can be made regarding the results in Figure 8. Firstly, it is noted that for small numbers of rays, the time consumption does not increase substantially as the number of rays increases. This is because the NVIDIA OptiX raytracing engine used for these simulations is well-optimized, and many ray paths are traced simultaneously. Consequently, the time consumption does not increase until the number of rays exceed some threshold which seems to be at about 10,000 rays in this case. This value is consistent across all algorithms.

Secondly, it is noted that the overall simulation time for a given number of rays is smaller in case B than in case A. This is caused by the increased absorption in case B. Each individual ray is propagated until its energy falls below some threshold, and the primary mechanism for energy reduction is reflection in an absorptive surface. When more absorption is introduced, the rays decrease in energy more quickly and are thus terminated more quickly. Once all rays have been terminated, the trace is complete.

This effect is also a probable explanation for the shorter simulation time for high numbers of rays in the diffuse field algorithm. In the diffuse field scattering algorithm, the energy of each ray is reduced according to the scattering parameter in addition to the energy reduction introduced by the absorption parameter. Although this energy is accounted for in the diffuse field model, it leads to a faster termination of the rays nonetheless.

Although the diffuse field algorithm is faster for high numbers of rays, the same is not true for small numbers of rays. This is likely due to the additional computational overhead associated with the additional raytracing step for calculating the diffuse field decay rate. This is the reason for the longer overall simulation time seen for the diffuse field algorithm in case A in Table 4.

These results indicate some differences between the algorithms. In test spaces where a relatively low number of rays are needed, the on–off and perturbation algorithms are faster. However, when more rays are needed the diffuse field scattering algorithm outperforms the others in terms of speed.

5.3. Usability

Finally, the usability of the three algorithms is studied, primarily with regards to the effects of changes to the scattering parameters. In Figure 9, the effects on room acoustic parameters as the scattering value changes are shown for the octave band centered on 1000 Hz. In case B especially, increases to the scattering parameter lead to decreased reverberation as measured both by T_{20} and EDT. It also leads to increased speech clarity C_{50}. The reduced reverberation is consistent with previous research in similar test spaces [10,13], while the previous results for speech clarity are inconclusive.

Changes to the ceiling absorbent scattering parameter do not affect the overall simulation outcome. This is explained by the high absorption coefficient of this surface. As shown in Table 1, the absorbent ceiling has an absorption coefficient of 0.95. Since scattering only acts on the energy which is reflected from a surface, even a scattering parameter of 1 can affect only 5% of the total incident energy on this surface. In this case, this seems to not be sufficient to have an impact on the simulated sound field. In the remainder of this section, only the scattering parameter for the remaining surfaces is considered.

From a technical perspective, the correlation between decreased reverberation and increased scattering in the on–off and perturbation scattering algorithms can be explained by considering the path of individual rays. If there is no scattering, a ray might remain bouncing between parallel walls for a long time. If these walls are relatively far apart, the energy of the ray decreases slowly. Towards the end of the simulation only these rays remain, and they thus dominate the late response. This can be seen in the simulation results as prolonged energy decay and consequently a long reverberation time. When on–off or perturbation scattering is introduced, the likelihood that a ray should be reflected away from that state increases and on average, each ray will be reflected between parallel walls for a shorter time. The overall reverberation is thus decreased.

In the diffuse field algorithm, the directions of individual rays are not affected by changes to the scattering parameter. Instead, increases to the diffuse field scattering parameter is directly interpreted as a transfer of energy (and thus an energy reduction) in each ray reflection in the traditional raytracing segment of the model. Therefore, the reverberation in this part of the model trivially decreases as the diffuse field scattering parameter increases. Furthermore, the energy transferred from the image source model in this way is added to the diffuse field, which has a decay rate corresponding to rays reflected in a random direction in each reflection. This matches simulations using the on–off scattering or perturbation algorithm with scattering parameters equal to the maximum value 1 and decay quickly, based on the argument in the previous paragraph. The reverberation found in the diffuse field scattering algorithm is thus doubly decreased by increases to the scattering parameter.

As seen in Figure 9, the changes to the sound field caused by changes in the scattering parameter are much more significant in case B than in case A. This is due to the difference in the distribution of absorption between the two cases. In case A, redirection of acoustic energy cannot cause it to immediately interact with a much more absorptive surface as none exist. However, in case B scattering effects will more often lead to reflection in a very absorptive surface and immediate energy reduction. Consequently, the effects of scattering in raytracers are much more pronounced in spaces with unevenly distributed absorption.

An aspect of usability is that the effects of changes to the scattering parameter should be predictable. The significant differences between case A and case B, and the two types of surface in case B, show that the effects of changes to the scattering parameters, to some extent, are unpredictable for all three algorithms. However, the perturbation algorithm is more unpredictable than either the on–off scattering algorithm or the diffuse field algorithm. In Figure 9, it is seen that all the room acoustic parameters may either increase or decrease as the perturbation scattering parameter increases. This makes the perturbation algorithm more difficult to use. In cases where the perturbation scattering parameter needs to be tuned, it may be difficult or impossible to determine whether it should be increased or decreased. Sometimes, both options may work and sometimes neither.

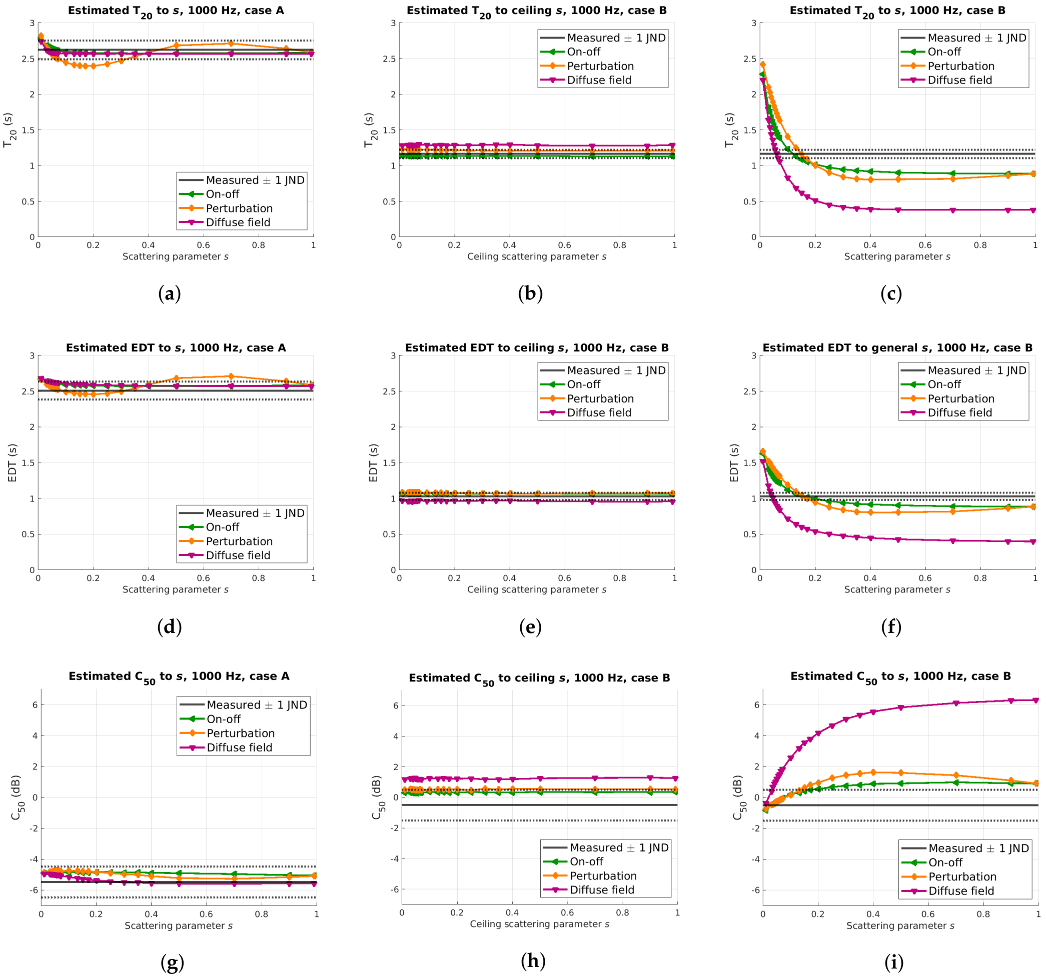


Figure 9. Simulated T_{20} (**a**–**c**), EDT (**d**–**f**) and C_{50} (**g**–**i**) as they change when the scattering parameter is modified. For the absorbing ceiling, changes to the scattering parameter does not affect the simulation results. In general, increasing the scattering parameter for the remaining surfaces leads to a decrease in reverberation (as shown in both T_{20} and EDT) for the on–off and diffuse field algorithms. It is also seen that the effects of the scattering parameter are much more significant in case B, for all algorithms but for the diffuse field algorithm in particular.

Another aspect of usability is whether small changes to the input parameter leads to small changes in the simulation output. As seen in Figure 9, the magnitude of the changes in simulated room acoustic parameters vary significantly depending on the space in question and the absolute value of the scattering parameter. In general, it seems that the effects of changing the scattering parameter is much smaller in spaces with more uniform absorption, which is an expected result. It also seems that modifications to the scattering parameters when it is fairly large has a relatively small impact on the simulated room acoustic parameters. These results replicate previous knowledge [10,13,28].

For the diffuse field algorithm in case B, the output variations are particularly large for small values in the scattering parameter (as shown in Figures 3 and 4). In some cases, minimal changes might lead to changes more prominent than 1 JND in the simulated room acoustic parameters. This shows that the diffuse field algorithm in some cases is very sensitive to variations in the scattering parameter, which consequently must be provided to a very high level of accuracy. Since the diffuse field scattering parameter does not have a well-defined physical counterpart, it cannot be measured and is very difficult to estimate, making it more or less impossible to estimate to the required level of accuracy.

Finally, an aspect of usability is whether reasonable input values lead to reasonable output values. In Section 3.3, it is determined that a fair initial guess of the scattering parameters at mid-frequency would be 0.1. A review of the results shown in Figure 9 shows that the results for this value are generally within a few JNDs for all algorithms in case A and the on–off and perturbation algorithms in case B. These results are not good but can be considered reasonable.

6. Discussion

Regarding physical accuracy, the diffuse field and perturbation algorithms perform worse than the on–off algorithm. However, their results show similar trends. They all show acceptable results for spatial averages in high frequencies, but cannot accurately model the sound field below 500 Hz, and do not emulate the spatial variations. The poor performance for low frequencies is consistent with the known limitations of raytracers [1,11], although it extends well above the Schröder frequency for the space in question. The discrepancies between measurements and simulations regarding spatial variations are somewhat unexpected. In any case, these issues are not useful to distinguish between the algorithms.

In general, the differences between the algorithms are seen much more clearly in the results for case B compared to case A. This suggests that surface scattering has a larger impact on the sound field in spaces with non-uniformly distributed absorption. Consequently, the choice of scattering algorithm and parameter requires extra care in spaces with localized absorption surfaces.

The diffuse field algorithm underperforms, as mentioned, in terms of physical accuracy in case B. The explanation presented in the analysis is that its underlying diffuse field model does not accurately describe a significant part of the acoustic field in case B. Since the model is inaccurate, the scattering algorithm based on it produces inaccurate simulation results. This explanation suggests that a more accurate physical model for the scattered energy can be developed and used as a foundation for a more accurate simulation algorithm, which still functions in a way similar to the diffuse field scattering algorithm. An example of such an extension is presented in [23], and it may also be possible to develop models based on acoustic radiosity or even Statistical energy analysis (SEA) similar to the model in [32]. The model must lend itself to fast and stable numerical calculations so that the benefits seen for the diffuse field algorithm in terms of time consumption are not lost.

The diffuse field scattering parameter and its effects on the simulation results are also affected by the inaccuracy of the diffuse field model in case B, negatively impacting the algorithm's usability. The diffuse field simulation results are exceedingly sensitive to variations in the diffuse field scattering parameter, and the tuned parameter values are different from what was expected. As mentioned in Section 2, it was mentioned that the diffuse field scattering parameter can be interpreted as a weighting coefficient, determining the relative influence of the underlying image source and diffuse field models. With this interpretation, the small value of the diffuse field scattering parameter in case B indicates that the diffuse field model is more inaccurate in describing the sound field in the space as compared to the image source model. In addition, the high sensitivity in simulation outcome shows that there are large differences between the two models in this case.

One of the differences between the on-off and perturbation scattering algorithm is that the perturbation scattering algorithm is less predictable in terms of changes to its scattering parameter. It was noted that increases in the perturbation scattering parameter might lead to increases or decreases in reverberation time. This study has not investigated whether this may be physically accurate, but it has been identified as a significant issue for the usability of the perturbation scattering algorithm.

The tuned scattering parameters are different for the three algorithms, showing that it is difficult or impossible to define a single value that would work for all algorithms. It was noted already in Section 2 that the physical interpretation of the scattering parameter differs in the various algorithms. It is consequently not surprising to see that different values are appropriate. However, it underscores the fact that caution should be exercised when setting up the material properties in a simulation, as table values appropriate for one algorithm may not be appropriate for a different one. This emphasizes the role of the simulation engineer and their experience with the simulation software in achieving accurate simulation results. In future research, measured surface scattering coefficients as defined in ISO 17497-1 and ISO 17497-2 [17,18] could be tried as scattering parameters in various algorithms. Doing so would improve understanding of the connection between the physical phenomena of surface scattering and how it is measured and the algorithms' scattering parameters.

This study has not indicated a significant difference between the three algorithms in terms of simulation time for the test space in question. The diffuse field algorithm seems able to reliably estimate the reverberation time in a simple space with relatively few rays, but this property does not extend to other room acoustic parameters. If many rays are used, the diffuse field algorithm is faster for the space examined in this study. It is unclear whether these results can be generalized. Notably, the diffuse field algorithm is prone to errors when the sound field deviates from the diffuse field approximation. Consequently, if a highly non-diffuse sound field causes the need for high numbers of rays, the diffuse field algorithm is inappropriate regardless of its fast calculation speed.

7. Conclusions

The goal of this study is to evaluate the influence of the scattering algorithm in acoustic raytracers. To this end, three algorithms are compared: The on–off scattering algorithm, the perturbation scattering algorithm and the diffuse field scattering algorithm. It is found that the choice of scattering algorithm impacts raytracers' physical accuracy, simulation speed, and usability. In addition, the scattering parameter must be chosen based on the scattering algorithm used.

Out of the on–off, perturbation and diffuse field scattering algorithms, the on–off scattering algorithm is most appropriate for raytracers based on its better usability and physical accuracy. A more complex test space with directional scatterers may yield further insights into the on–off scattering algorithm's strengths and weaknesses.

Author Contributions: Conceptualization, H.A., N.-G.V. and D.B.H.; methodology, H.A.; software, H.A.; formal analysis, H.A.; investigation, H.A. and N.-G.V.; writing—original draft preparation, H.A.; writing—review and editing, H.A., N.-G.V. and D.B.H.; visualization, H.A.; supervision, N.-G.V. and D.B.H.; funding acquisition, D.B.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Swedish Research Council grant number 2016-01784. The APC was partially funded by Lund University.

Data Availability Statement: Raw data and source code are available upon request.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Kuttruff, H. *Room Acoustics*, 6th ed.; CRC Press: Boca Raton, FL, USA, 2017.
2. Vorländer, M. *Auralization. Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*; Springer: Berlin, Germany; 2008; pp. 1–335. [CrossRef]
3. Krokstad, A.; Strom, S.; Sørsdal, S. Calculating the acoustical room response by the use of a ray tracing technique. *J. Sound Vib.* **1968**, *8*, 118–125. [CrossRef]
4. Dalenbäck, B.I.; Mendel, K.; Svensson, P. A Macroscopic View of Diffuse Reflection; Audio Engineering Society, Inc.: New York, NY, USA, 1993.
5. Lam, Y.W. The dependence of diffusion parameters in a room acoustics prediction model on auditorium sizes and shapes. *J. Acoust. Soc. Am.* **1996**, *100*, 2193–2203. [CrossRef]
6. Hodgson, M. Evidence of diffuse surface reflections in rooms. *J. Acoust. Soc. Am.* **1991**, *89*, 765–771. [CrossRef]

7. Naylor, G.; Rindel, J. Predicting room acoustical behavior with the ODEON computer model. *Acoust. Soc. Am.* **1992**, *92*, 2346–2346. [CrossRef]
8. Dalenbäck, Bengt-Inge. CATT-Acoustic v9.1 Powered by TUCT v2. 2020. Available online: https://www.catt.se/ (accessed on 28 August 2021).
9. Farina, A. RAMSETE-a new Pyramid Tracer for medium and large scale acoustic problems. In Proceedings of the Euronoise 95 Conference, Lyon, France, 21–23 March 1995.
10. Schröder, D. Physically Based Real-Time Auralization of Interactive Virtual Environments. Ph.D. Thesis, Institute of Technical Acoustics, Aachen, Germany, 2011.
11. Savioja, L.; Svensson, P. Overview of geometrical room acoustic modeling techniques. *J. Acoust. Soc. Am.* **2015**, *138*, 708–730. [CrossRef] [PubMed]
12. Cox, T.J.; D'Antonio, P. *Acoustic Absorbers and Diffusers*, 3rd ed.; Taylor & Francis Group: Boca Raton, FL, USA, 2017.
13. Zhu, X.; Kang, J.; Ma, H. The impact of surface scattering on reverberation time in differently shaped spaces. *Appl. Sci.* **2020**, *10*, 4880. [CrossRef]
14. Shtrepi, L.; Astolfi, A.; Pelzer, S.; Vitale, R.; Rychtáriková, M. Objective and perceptual assessment of the scattered sound field in a simulated concert hall. *J. Acoust. Soc. Am.* **2015**, *138*, 1485–1497. [CrossRef] [PubMed]
15. Kuttruff, H. *Room Acoustics*, 5th ed.; CRC Press: Boca Raton, FL, USA, 2009.
16. Morse, P.M.; Ingard, K.U. *Theoretical Acoustics*; McGraw-Hill Book Company: New York, NY, USA, 1968.
17. ISO 17497-2:2012. *Acoustics–Sound-Scattering Properties of Surfaces–Part 2: Measurement of the Directional Diffusion Coefficient in a Free Field*; Technical Report; ISO: Geneva, Switzerland, 2012.
18. ISO 17497-1:2004. *Acoustics–Sound-Scattering Properties of Surfaces–Part 1: Measurement of the Random-Incidence Scattering Coefficient in a Reverberation Room*; Technical Report; ISO: Geneva, Switzerland, 2004.
19. Cox, T.J.; Dalenback, B.I.; D'Antonio, P.; Embrechts, J.J.; Jeon, J.Y.; Mommertz, E.; Vorländer, M. A tutorial on scattering and diffusion coefficients for room acoustic surfaces. *Acta Acust. United Acust.* **2006**, *92*, 1–15.
20. Pilch, A. Optimization-based method for the calibration of geometrical acoustic models. *Appl. Acoust.* **2020**, *170*. [CrossRef]
21. Postma, B.N.J.; Katz, B.F.G. Creation and calibration method of acoustical models for historic virtual reality auralizations. *Virtual Real.* **2015**, *19*, 161–180. [CrossRef]
22. Vorländer, M. Computer simulations in room acoustics: Concepts and uncertainties. *J. Acoust. Soc. Am.* **2013**, *133*, 1203–1213. [CrossRef] [PubMed]
23. Lam, Y.W. A comparison of three diffuse reflection modeling methods used in room acoustics computer models. *J. Acoust. Soc. Am.* **1996**, *100*. [CrossRef]
24. ISO 3382-1:2009. *Acoustics–Measurement of Room Acoustic Parameters–Part 1: Performance Spaces*; Technical Report; ISO: Geneva, Switzerland, 2009.
25. ISO 354:2003. *Acoustics–Measurement of Sound Absorption in a Reverberation Room*; Technical Report; ISO: Geneva, Switzerland, 2003.
26. Meissner, M. Acoustics of small rectangular rooms: Analytical and numerical determination of reverberation parameters. *Appl. Acoust.* **2017**, *120*, 111–119. [CrossRef]
27. ODEON Room Acoustics Software. 2020. Available online: https://odeon.dk/ (accessed on 15 September 2021).
28. Dalenbäck, B.I. Reverberation Time, Diffuse Reflection, Sabine, and Computerized Prediction. 1999. Available online: https://www.catt.se/diffseries/index.htm (accessed on 21 September 2021).
29. Berzborn, M.; Bomhardt, R.; Klein, J.; Jan-Gerrit, R.; Vorländer, M. The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic Measurements and Signal Processing. In Proceedings of the 43rd Annual German Congress on Acoustics, Kiel, Germany, 6–9 March 2017.
30. Parker, S.G.; Bigler, J.; Dietrich, A.; Friedrich, H.; Hoberock, J.; Luebke, D.; McAllister, D.; McGuire, M.; Morley, K.; Robison, A.; et al. OptiX: A general purpose ray tracing engine. *ACM Trans. Graph.* **2010**, *29*. [CrossRef]
31. Odeon A/S. ODEON Room Acoustics Software–User's Manual.2020. Available online: https://odeon.dk/download/Version16/OdeonManual.pdf (accessed on 15 September 2021).
32. Nilsson, E. Decay Processes in Rooms with Non-Diffuse Sound Fields Part II : Effect of Irregularities. *Building* **2004**, *11*, 133–143. [CrossRef]

Paper C

acoustics

IMPACT FACTOR 2.1

CITESCORE 4.0

Article

A Novel Algorithm for Directional Scattering in Acoustic Ray Tracers

Hanna Autio and Erling Nilsson

MDPI

https://doi.org/10.3390/acoustics5040054

acoustics

Article

A Novel Algorithm for Directional Scattering in Acoustic Ray Tracers

Hanna Autio [1,*] **and Erling Nilsson** [2]

[1] Division of Engineering Acoustics, Lund University, P.O. Box 118, 221 00 Lund, Sweden
[2] E. Nilsson Akustik, Flygelvägen 275, 224 72 Lund, Sweden; erni0420@gmail.com
* Correspondence: hanna.autio@construction.lth.se

Abstract: It is vital to consider acoustic scattering when using geometrical acoustic simulation techniques, such as ray tracing. However, there are few methods for modelling scattering, and most rely on strong assumptions of uniformity on the distribution of scattered energy. In this paper, a model for directional scattering in ray tracers is presented. The model is based on an idealized model of a 1D scatterer, which is then used to extend the most commonly used scattering algorithm in ray tracers today. The developed algorithm is implemented in a ray tracer and tested to evaluate its performance compared to existing methods. It is found that the directional scattering algorithm can be used to replicate measured effects on room acoustic parameters caused by changes in the orientation of 1D scatterers.

Keywords: acoustic ray tracing; scattering algorithm; 1D scatterer; acoustic scattering; 1D diffuser; diffuser modelling; room acoustic modelling

1. Introduction

Geometrical acoustics is a class of commonly used methods for acoustic simulation [1–4] where a simplified model of sound propagation is used. One of the most popular models in geometrical acoustics is ray tracing, as it is fairly computationally efficient, accurate at high frequencies and relatively intuitive to implement and use. An important element of acoustic ray tracing is the incorporation of scattering, which has been shown to improve the accuracy of simulation software [5]. In this paper, a novel ray-tracing scattering algorithm is developed. It extends the most commonly used algorithm to account for directional variations in the scattering pattern.

Although there are several application of acoustic ray tracing, the ray tracer used in this project is developed specifically for room acoustic modelling. The desired accuracy of room acoustic modelling is limited by the limits of human perception. The ray tracer used in this project was developed in-house and models the sound field as energetic particles that travel along straight paths between the surfaces of an enclosure. The phase of the particles is not recorded, and their energy is assumed to be affected only by air attenuation and surface absorption at the reflections.

There are several possible definitions of acoustic scattering. In this project, the definition suggested by Morse and Ingaard [6] is used. Scattering is thus defined as the difference between the true reflection and the idealized plane wave reflection on an ideal surface. In the context of ray tracing, the ideal plane wave reflection is the mirror reflection of the acoustic ray [7]. The distribution of the scattered energy is most commonly modelled as diffuse, even in cases where the real distribution is expected to have significant directional components [3,8].

Models for surface scattering and in particular scattering due to diffraction represent a major research field within acoustics [9–12]. Several models for the scattering of acoustic energy have recently been presented for use in many different models [13–17]. In particular,

check for updates

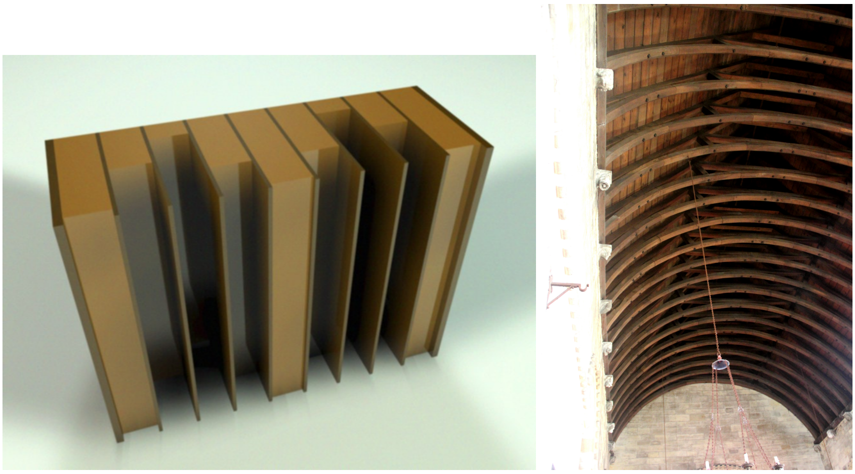
Citation: Autio, H.; Nilsson, E. A Novel Algorithm for Directional Scattering in Acoustic Ray Tracers. *Acoustics* **2023**, *5*, 928–947. https://doi.org/10.3390/acoustics5040054

Academic Editor: Michel Darmon

Received: 11 September 2023
Revised: 15 October 2023
Accepted: 19 October 2023
Published: 24 October 2023

methods to include scattering from diffraction have been introduced in multiple tools for geometric modelling [1,18,19]. The approach to scattering suggested in this paper is not limited to diffraction, and it is aimed at providing an approximate method comparable to the methods for uniform surface scattering that are frequently implemented in ray tracers today.

Structural elements that scatter acoustic energy predominantly in one direction may occur incidentally from construction elements such as battened ceilings or stairs but may also be installed with acoustic intent to redirect acoustic energy or create a more diffuse sound field [3,20]. These types of structures are referred to as 1D scatterers in this paper, and some examples are shown in Figure 1. In spaces with very localized absorption, for example, public spaces with ceiling absorbers, it has been shown that the orientation of 1D scatterers can have a significant impact on the acoustic environment [21,22]. This indicates that there is a need to model the directional variations of such structures.



(**a**) (**b**)

Figure 1. Examples of 1D scatterers. (**a**) One-dimensional acoustic diffuser. (Source: Musikproducer, CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0 (accessed on 18 October 2023), via Wikimedia Commons). (**b**) Truss on church ceiling. (Source: Storye book, CC BY 4.0 https://creativecommons.org/licenses/by/4.0 (accessed on 18 October 2023), via Wikimedia Commons).

In acoustic ray tracing, there are two traditional options to model 1D scatterers. Either they are modelled geometrically, preserving the directional variations, or they are modelled as flat surfaces with scattering coefficients. The flat-surface model with scattering coefficients is easy to adapt to different frequency ranges, but all directional variations are lost with the standard scattering algorithms. Using the geometric modelling strategy works well at high frequencies and can accurately replicate the directional variations but risks overestimating the scattering effects of the structure at low frequencies.

Sound waves interact with structural variations in different ways, depending on the relative size of the structural variations to the acoustic wavelength. When the wavelength of an impeding acoustic wave is much larger than the structural variations on a surface, acoustic energy is reflected with very little scattering. When the wavelengths are of approximately the same size as the surface structures, the reflected acoustic energy is affected by diffraction and temporal dispersion and thus scattered in a complicated pattern. For wavelengths much smaller than the surface variations, acoustic rays react more locally and are reflected in the surface segment they hit [23]. In room acoustic simulation, the modelled wavelength varies between about 3.5 m and 4 cm, and an appropriate spatial resolution for the geometric model accordingly varies between about 1 m and 1 cm. Generally, however, only one model of the geometry is developed and used for all wavelengths. If the model is detailed and the resolution high, it works well at high frequencies but may overestimate the scattering effects at low frequencies. If the model has low spatial resolution, it works better at low frequencies, and higher scattering parameters can be assigned to surfaces with structural variations smaller than the model resolution to compensate for the reduced resolution. Typically, only uniform scattering is modelled using the scattering parameters. In order to preserve possible directional variations in the simplified surfaces, a directional scattering algorithm would be useful.

It has been shown that the direction of scattering can affect room acoustic parameters in some cases [20–22]. In these papers, the orientation of 1D scatterers was found to have a perceivable effect on several room acoustic parameters. This indicates the need for an algorithm that can account for directional variations in acoustic scattering and does not rely on a high spatial resolution that may adversely affect the simulation performance at low frequencies or in terms of computational speed.

The algorithm introduced in this paper extends the commonly used uniform model of ray scattering by introducing a directional variation in the distribution of the scattered energy. It is capable of emulating the effects on room acoustic parameters caused by changes in the orientation of 1D scatterers and can easily be adjusted across the full frequency range.

2. A Directional Scattering Algorithm

In this section, a novel directional scattering algorithm for acoustic ray tracers is developed. An initial model is developed based on an idealized 1D scatterer and used to extend the most common scattering algorithm used in ray tracers today. Its implementation is described, and some comments are made regarding parameter selection.

2.1. Model for an Ideal 1D Scatterer

In order to determine a suitable scattering model for a 1D scatterer, an idealized version is considered. Consider a surface such as the one illustrated in Figure 2. A local orthonormal coordinate system is defined, such that the x-axis coincides with the surface normal and the y-axis falls along the direction of maximum scattering. For the scatterer shown in Figure 2, the y-axis is thus perpendicular to the periodic surface elements and the z-axis parallel to them. In the idealized case, the acoustic energy reflected from the surface is diffusely distributed along the y-axis, while the directional information is preserved along the z-axis. A sample from the resulting distribution of possible reflections for a single ray is shown in Figures 3 and 4. The reflected energy is distributed in a semicircle, proportional to the cosine of the angle to the semicircle's centre. All energy is reflected specularly in the z-dimension, thus preserving the directionality of the sound field in this dimension.

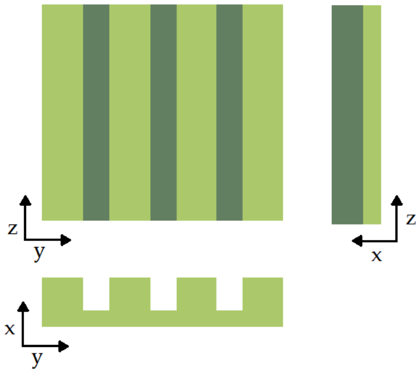
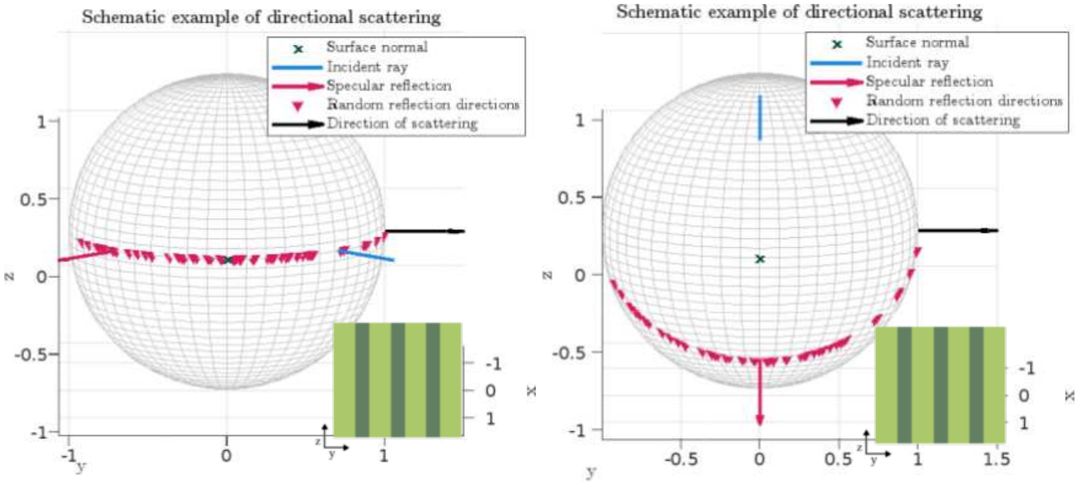


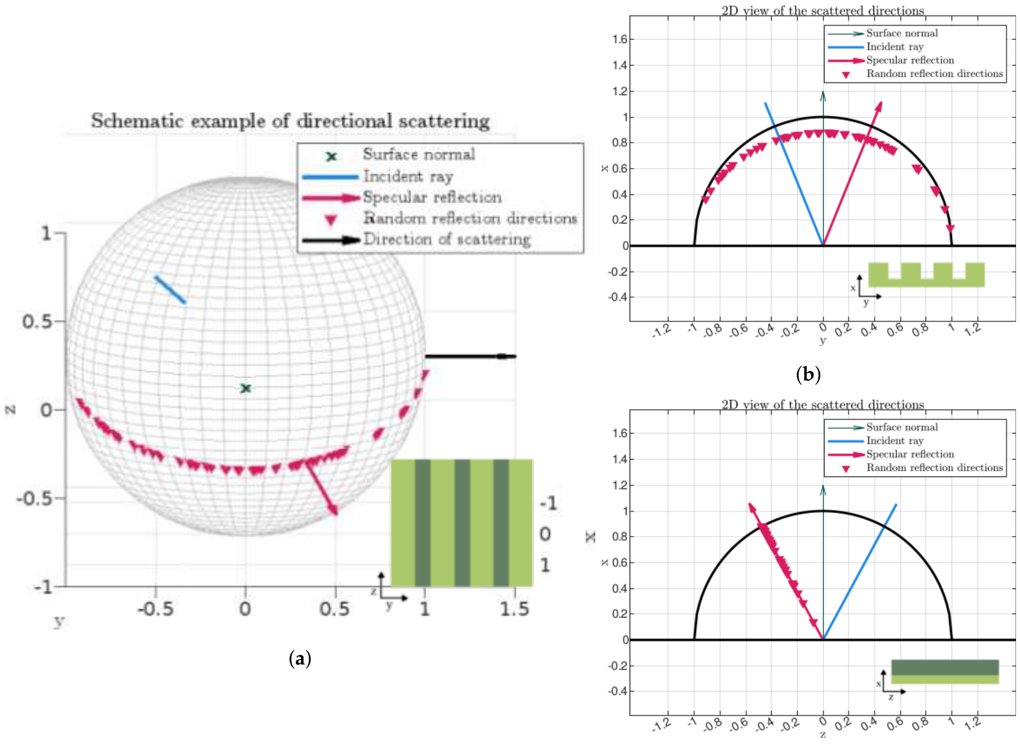
Figure 2. Simplified illustration of a 1D scatterer. Acoustic energy is dispersed in the y-direction, but directional information is retained along the z-axis.



(**a**)

(**b**)

Figure 3. Illustration of possible reflection directions for an idealized 1D scatterer for incident rays parallel (**a**) and perpendicular (**b**) to the direction of scattering. Insets illustrate the orientation of the scatterer.



(**a**)

(**b**)

(**c**)

Figure 4. Illustrations of possible reflection directions from the idealized 1D scatterer. (**a**) gives an overview. In (**b**), it is shown that the ray is reflected specularly along the z-axis, and in (**c**), the diffuse nature of reflection along the y-axis is illustrated. Insets illustrate the orientation of the scatterer.

This model can be further understood by considering which information is preserved and which is destroyed in the reflection. The direction of a reflected ray is completely defined by two coordinates, for example, its azimuth and elevation on the unit sphere around the point of reflection. When the ray is reflected specularly, the two coordinates are completely determined by the direction of incidence and the normal. When the ray is reflected in a random direction, both coordinates are sampled from some random distribution. In the model defined above, one random number is generated to ensure that the energy is diffusely distributed on the semicircle, whereas the remaining coordinate is determined based on the direction of the incident ray. In this way, the ray can be said to be partially scattered.

The model formulated above can be used to define a method for ray reflection in an ideal 1D scatterer, but further development is necessary if the algorithm is to be useful for more realistic problems. In such cases, some acoustic energy may be reflected altogether specularly or scattered in all dimensions. For such surfaces, an extension of the most common scattering algorithm is suggested. The most common method for random scattering used in ray tracers today is an algorithm where there are two possible outcomes when a ray is reflected by a surface [3,7,23]. Either the ray is reflected specularly, in the mirror direction, or the ray is scattered and reflected in a random direction sampled from an ideally diffuse energy distribution. To account for directional scattering, a third option is introduced wherein the ray is reflected as in the case of an ideal 1D scatterer.

2.2. Implementation

In the case of uniform surface scattering, the scattering properties of the surface are described by a single frequency-dependent scattering coefficient, s. This value is used to randomly determine, for each incident ray, whether it is scattered or reflected specularly, and the coefficient corresponds to the ratio of incident energy that is reflected away from the specular direction [5,24].

In the proposed directional scattering algorithm, the surface is described by two scattering coefficients, $0 \leq s_d \leq s_s \leq 1$, and a vector, $\mathbf{d}$, defining the direction of maximum scattering. In the idealized scatterer shown in Figure 2, $\mathbf{d}$ corresponds to basis vector $\mathbf{y}$. The directional scattering coefficients define the ratio of acoustic energy that is reflected according to different distributions. s_s defines the ratio of acoustic energy that is scattered. According to the definition used in this paper, this is all acoustic energy that is not reflected specularly. s_d defines the ratio of acoustic energy that is reflected diffusely, or uniformly. If $s_d = s_s$, there are no directional variations in reflected energy; all scattered energy is assumed to be diffusely distributed; and the directional scattering model coincides with the most common uniform scattering model. When $s_d < s_s$, some energy that is not diffusely reflected is scattered. In the directional scattering model, this energy is assumed to be distributed according to the model presented for the ideal 1D scatterer above.

These coefficients are used similarly to the single scattering coefficient in the uniform scattering algorithm to determine the type of reflection each ray undergoes. Pseudo code is presented in Listing 1, where a local coordinate system is used, consistent with what was used in Section 2.1. This is performed without loss of generality.

Listing 1. Pseudo code for the directional reflection algorithm.

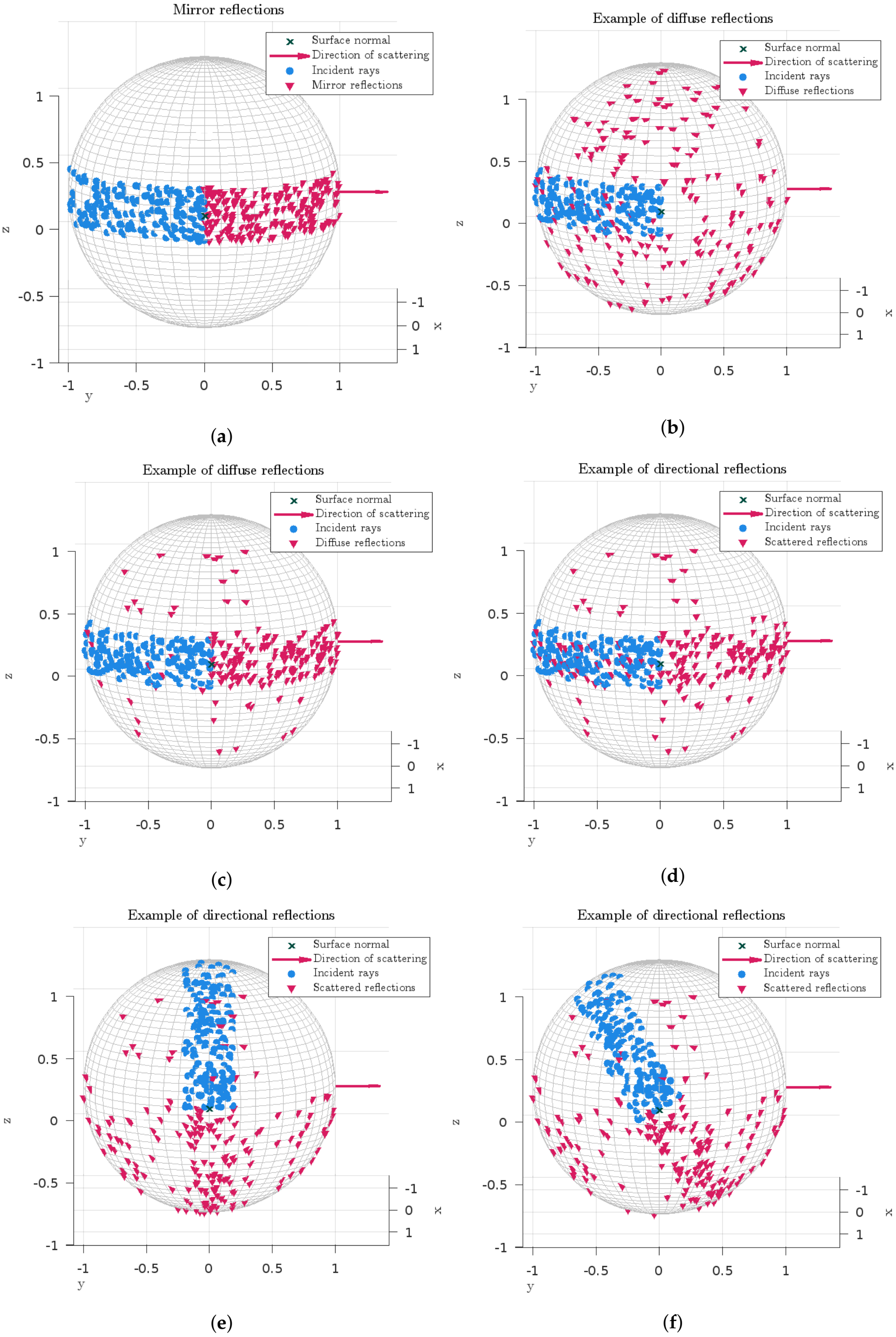
```
// Generate a random value to determine if the ray is scattered
float rnd=rand(0,1);

if (rnd < s_d) { // Ray is diffusely reflected
    new_direction=generate_lambert();
} else if (rnd > s_s) { // Ray is not scattered
    new_direction=reflect(incident, surface_normal);
} else { // Ray is partially scattered
    // Generate a random direction on a semi-circle
    new_direction=generate_cosine_circle();

    // Determine the desired angle between the xy-plane
    // and the reflected ray
    float z_angle=rotation_angle(xy_plane, incident_local);

    // Calculate the new direction by rotation around the
    // y-axis
    new_direction=rotate(new_direction, z_angle, y_axis);
}
```

In the directional scattering algorithm, a random value $0 \leq \texttt{rnd} \leq 1$ is firstly generated from a uniform distribution. This is used to determine the fate of the ray. If $\texttt{rnd} < s_d$, the ray is fully scattered, and the reflection direction is randomly sampled from a Lambertian 2D distribution. An example of ray directions generated from the Lambertian distribution is shown in Figure 5b.

If random value $\texttt{rnd} > s_s$, the ray is specularly reflected in the surface normal. An example is shown in Figure 5a.

If the ray is neither diffusely nor specularly reflected, it is partially scattered according to the model presented above. The reflection direction is generated in two steps, one random step and one deterministic step. A direction on the semicircle is randomly sampled from a cosine distribution. An example of possible directions in this step is shown in Figure 3a. The direction generated in this way is then rotated about the y-axis to ensure that the directional information in the z-dimension is preserved (see Figure 4b). The angle of rotation is determined as the angle between the incident ray and the xy-plane, and the final direction is generated using Givens rotation [25].

Some samples of possible ray reflections are shown in Figure 5. The incident ray directions are sampled from a subset of a diffuse distribution. This should indicate how the directional characteristics of the incident and reflected sound fields depend on the scattering algorithm used. In Figure 5a, there is no scattering, and the rays are specularly reflected. In Figure 5b, a fully diffused reflection model is used, and all directional characteristics of the sound field are lost. In Figure 5c, standard uniform scattering is used, and the reflected rays are either mirror reflections or fully random diffuse reflections as determined by the scattering parameter (s). In the last three plots (Figure 5d–f), the directional scattering algorithm is used. It can be seen that the acoustic energy is more dispersed in the direction of scattering. In Figure 5e, it can be seen that most rays continue to travel in the negative z-direction, thus maintaining a flow of acoustic energy in that direction.



(**a**) (**b**) (**c**) (**d**) (**e**) (**f**)

Figure 5. Illustration of some possible reflection patterns using the directional scattering algorithm. In (**a**), there is no scattering ($s_s = s_d = 0$), and all rays are specularly reflected. In (**b**), fully diffused reflections are shown ($s_s = s_d = 1$). In (**c**), the two scattering coefficients are equal to 0.2. This coincides with the standard uniform scattering algorithm with $s = 0.2$. In (**d**–**f**), examples of directional scattering are shown for changing direction of incidence. Here, $s_s = 0.6$, and $s_d = 0.2$.

2.3. Directional Scattering Coefficients

A central aspect of implementing the directional scattering algorithm is to determine appropriate scattering coefficients s_s, s_d, which in turn requires some more discussion on ray-tracing algorithms in general. According to the ray-tracing model, the ray after reflection is a sample from some more complex underlying distribution of acoustic energy after reflection. The sampling method and the model for the energy distribution are closely linked. Thus, the directional scattering algorithm developed in this paper defines a model for the distribution of acoustic energy after reflection. According to this model, the reflected energy consists of a superposition of a diffuse cloud, a disc or semicircle (as shown in Figures 3 and 4) and the mirror reflection. Scattering parameters s_s, s_d define the ratio of energy in each of these energy clouds. Before it can be determined which coefficients are realistic, it should be examined whether the overall model for the distribution of reflected energy is appropriate. This can be performed, for example, using numerical simulations of the wave equation or measurements of the full polar response from real 1D scatterers.

However, due to the similarity between the developed model and the widely used model for uniform scattering, it is possible to use experience with the latter to estimate which may be reasonable values for the scattering parameters. In the classical model, there is a single scattering parameter s. This scattering parameter s is similar to s_d in the directional model in that it defines the ratio of energy that is diffusely reflected. It is also similar to s_s in the directional model, as it defines the ratio of energy that is not reflected in the mirror direction, corresponding to the random-incidence scattering coefficient as described in the international standards [5,24]. While there are some table values for s, it is frequently estimated based on the surface properties and the size of surface irregularities [18,26].

For the purpose of simulation, it is suggested that scattering parameters s_s, s_d are estimated similarly to how s would be estimated, based on the geometry of the surface. s_s should be estimated by finding the direction of scattering (that is, the direction along which most acoustic energy is scattered) and then estimated along this direction based on the geometric variations found in that direction. s_d should be estimated using the same rationale but along the direction perpendicular to the direction of scattering.

3. Materials and Methods

The directional scattering algorithm developed in this study was evaluated through comparison with measurements and commonly used simulation strategies.

3.1. Measurements

Measurements were taken from an earlier study, presented and analysed in [20–22,27]. For a comprehensive analysis of the measured impulse responses, the reader is referred to those papers. In this study, a subset of the measurements was used and is briefly introduced in this paper.

Impulse response measurements were carried out using the DIRAC system (DIRAC type 7841, v.6.0). The sound source used was a dodecahedral loudspeaker positioned 1.55 m from the floor emitting an exponential sweep. The receiver was placed 1.2 m from the floor. A total of two source positions and six listener positions were measured, as indicated in Figure 6a. During the measurements used in this project, an absorbent ceiling was mounted 2.7 m from the floor, and there was no furniture.

Measurements were performed using 1D scatterers mounted on two walls, oriented to scatter acoustic energy horizontally (as shown in Figure 6b) or vertically (as shown in Figure 6c). Measurements were carried out using 0, 4, 8, 12, 16, 20 or 24 scatterers, either all oriented horizontally or all oriented vertically. When 8 and 12 scatterers were used, two different placement patterns were used but no significant differences between the two patterns could be observed.

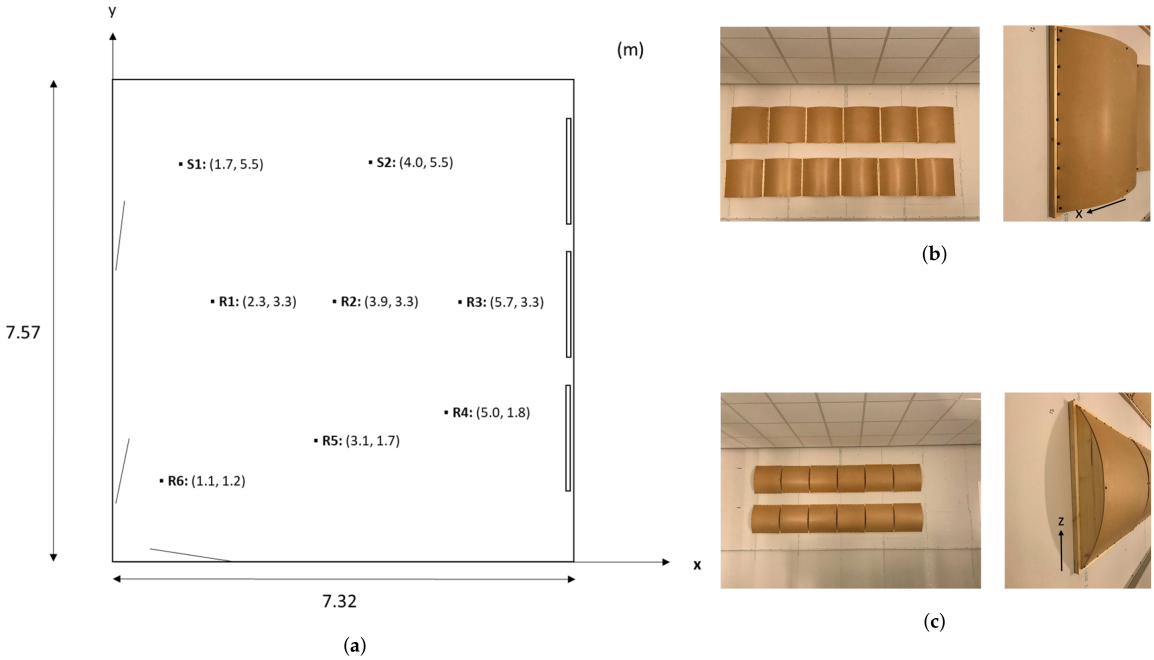


Figure 6. Figure (**a**) shows a sketch of the measurement space, and (**b**,**c**) show the scatterers used, oriented to scatter sound horizontally (in (**b**)) or vertically (as in (**c**)). The scatterers were mounted along the walls shown top and left in (**a**). An absorbent ceiling was mounted 2.7 m above the floor [20].

Room acoustic parameters were extracted from the impulse response measurements using `MATLAB`. The chosen parameters were reverberation time (T_{20}), early decay time (EDT) and speech clarity (C_{50}). While all tested room acoustic parameters were impacted by the orientation of the scatterers, the reverberation time (T_{20}) was used to evaluate the simulation results in most cases, spatially averaged in accordance with ISO standard 3382 [28].

3.2. Modelling the Scatterers in the Ray Tracer

A digital model of the measurement space was constructed and used for simulations. The base geometric model is shown in Figure 7a. The scatterers were modelled either using a geometric model (as shown in Figure 7b) or as flat surfaces with appropriate scattering and absorption coefficients.

Five sets of simulations were performed, with each modelling the scatterers in different ways. The different configurations are referred to as *Geo*, *Absorption only*, *Uniform scattering*, *Directional scattering A* and *Directional scattering B*. All configurations are more thoroughly described and motivated below. For each set of simulations, each measurement situation (in terms of scatterer position and orientation, and source and receiver positions) was simulated.

The Geo model is based on a geometric model for the scatterers, as shown in Figure 7b. This model is considered one of the baseline methods for modelling 1D scatterers, and the results from this model should be compared to the directional scattering model. The absorption coefficients for the scatterers were obtained from measurements, and the scattering coefficients (using the standard uniform scattering model) were estimated by what yielded accurate simulation results.

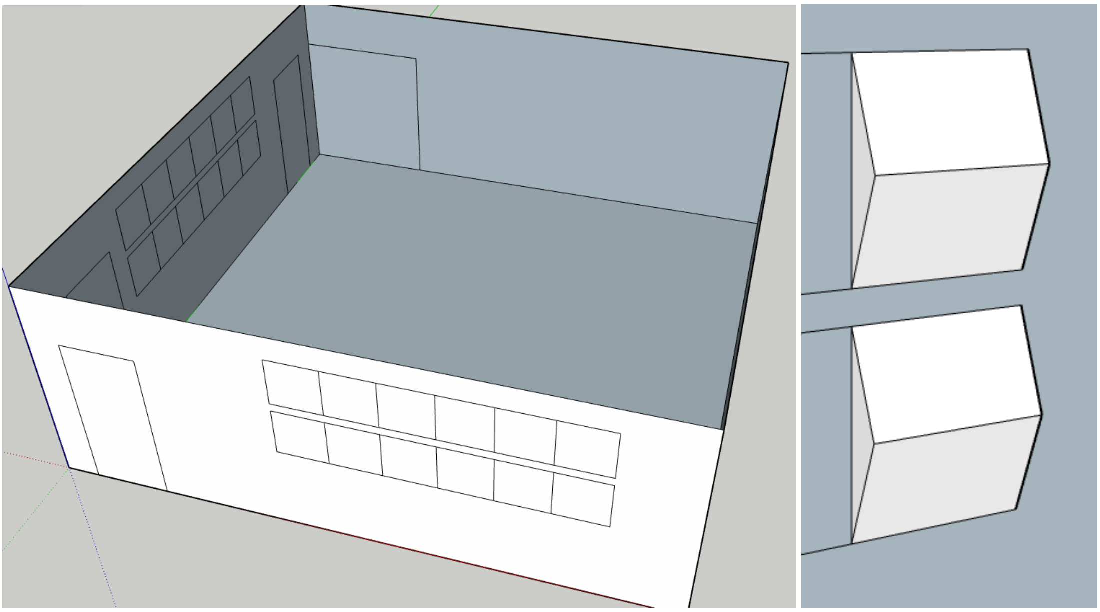

(a) (b)

Figure 7. The digital model used for simulation of the measurement space. Compare to Figure 6. In (**a**), the model for the whole space is shown with the possible positions of scatterers marked, and in (**b**), the geometric model of the scatterers is shown.

In the Absorption only model, the scatterers are modelled as flat surfaces with absorption coefficients based on measurements. Scattering coefficients are identical to those of the surrounding walls. This model is used to isolate the absorption effects introduced by the scatterers, so that the effects of scattering can be better identified and analysed.

The Uniform scattering configuration models the scatterers as flat surfaces with uniform scattering. The absorption coefficients are based on measurements, and the scattering coefficients are estimated based on rough measurements, as described below. This model is considered a second baseline method for modelling 1D scatterers, and the results should be thoroughly compared to measurements and the results from the directional scattering algorithm.

Directional scattering A and Directional scattering B implement the directional scattering algorithm presented in Section 2.2. The scatterers are modelled as flat surfaces with absorption coefficients based on measurements, a vector **d** defining the direction of scattering and two separate scattering coefficients. The vector **d** is vertical for the vertically oriented scatterers and horizontal for the horizontally oriented scatterers. The difference between configurations A and B lies in the scattering coefficients, which, in A, are estimated from measurements and, in B, are refined based on the results of simulations. A range of possible values should be examined to help determine an appropriate range of scattering coefficients for the directional algorithm.

3.3. Surface Absorption and Scattering Parameter Selection

Surface absorption and scattering parameters were obtained from table values, from measurements and by using an optimization algorithm. The floors and doors were found to have a low relative absorption area and to have a limited impact on the simulation result, and table values were used. The absorption properties of the scatterers were available from measurements [20], and the scattering parameters (in the Geo model) were estimated based on shape of the scatterers. The absorption and scattering parameters for the remaining surfaces, absorptive ceiling and walls, were estimated using an optimization procedure.

An iterative optimization method was used to simultaneously estimate the scattering and absorption parameters of the ceiling and the walls. A geometric strategy (Geo) was used to model the scatterers in this step. In order to ensure that the estimated parameters would work for different scatterer configurations, measurement setups using 0, 12 or 24 scatterers mounted horizontally or vertically were used. A least squares cost function was set up to calculate the error of the simulated reverberation time (T_{20}) compared to the measurements. An initial range of possible absorption and scattering values were estimated from table values and adjusted so that the configuration with minimal absorption resulted in an overestimation of the reverberation time and the maximal absorption resulted in an underestimation. The upper and lower bounds were used to define a grid in the parameter space over which the different scatterer configurations were simulated. After an initial set of simulations, the optimal setup was used as a basis for an additional step with a more refined grid. The resulting parameters are seen in Tables 1 and 2, and some results of the optimization are shown in Section 4.2.

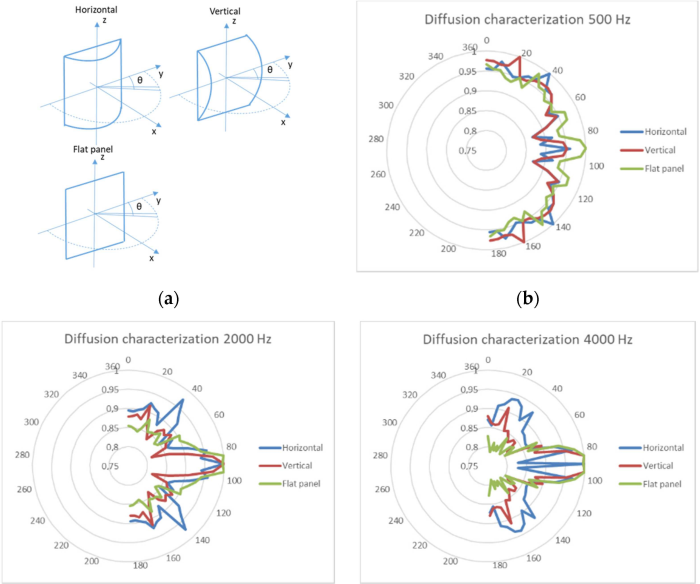
Table 1. Absorption parameters α used for simulations.

f (Hz)	Floor	Doors	Scatterers	Walls	Absorptive Ceiling
125	0.01	0.2	0.72	0.13	0.21
250	0.01	0.2	0.48	0.15	0.17
500	0.01	0.04	0.49	0.07	0.28
1000	0.01	0.02	0.32	0.04	0.60
2000	0.01	0.02	0.18	0.07	0.50
4000	0.01	0.02	0.02	0.06	0.80

Finally, the scattering parameters used in the directional scattering models were determined. As the algorithm is new, there is no precedent on how the parameters should be estimated. In this paper, an approach is suggested where scattering parameter s_s is assumed to be related to the scattering effects in the direction of most scattering, and s_d is related to the amount of scattering in the direction perpendicular to this. With this interpretation, s_s for the scatterers used in this project can be estimated by measuring the scattering effects on the horizontal plane about a horizontally oriented scatterer, and s_d, from measurements on the vertical plane around the same scatterer.

Using this model, an initial guess was estimated from single-dimensional measurements of the diffusion characteristics of the scatterers. The measurements are shown in Figure 8. They were performed on a single scatterer with a surface area of 60×60 cm^2, and edge effects were expected to have a significant effect on the measurements.

Estimates of scattering parameters were obtained from the measured diffusion characteristics. The diffusion characteristics as shown in Figure 8 were estimated by measuring the reflected power in 5-degree intervals on the azimuth plane around the scatterers for both orientations. In addition, a flat board of the same dimensions was measured in the same way. A single-dimensional scattering coefficient was estimated from the measurements, based on the model presented in the standards [5,24]. The coefficients obtained this way overestimate the true scattering effects of the scatterers, as they also account for the edge effects. An attempt to correct for this was made by using the differences in the estimated coefficients. It was assumed that the flat board could accurately be modelled using the scattering coefficients used for the walls. The difference in the scattering coefficient between flat board and horizontally oriented scatterer was assumed to match the difference in scattering parameter between a flat wall and the scattering coefficient in the direction of scattering, s_s. Scattering parameter s_d, related to the direction parallel to the ridge of the scatterer, was estimated in the same way but using measurements from the vertical scatterer. The scattering parameters obtained in this way were used for the Directional scattering A simulation model. In addition, the average between s_s and s_d was used to approximate the scattering parameters in the Uniform model.



(**a**) (**b**)

(**c**) (**d**)

Figure 8. Measured diffusion characteristics of the scatterers used in this project. Picture obtained from [22] with permission. In (**a**), the orientation of the scatterers are shown the different measurements. The diffusion characteristics are shown for each orientation in (**b**–**d**) for frequencies 500 Hz, 2000 Hz and 4000 Hz respectively.

As previously noted, there is no precedent for the estimation method or interpretation of the scattering parameters used in the directional scattering model. Initial simulations indicated that the differences between s_s and s_d needed to be quite large to cause a noticeable effect due to the orientation of the scatterers, significantly larger than what was found for Directional scattering A. In the Directional scattering B model, the scattering parameters were determined by what yielded a reasonable simulation result.

Table 2. Scattering parameters used in simulations. Except where otherwise noted, the parameter used is the single parameter used in the common uniform scattering model.

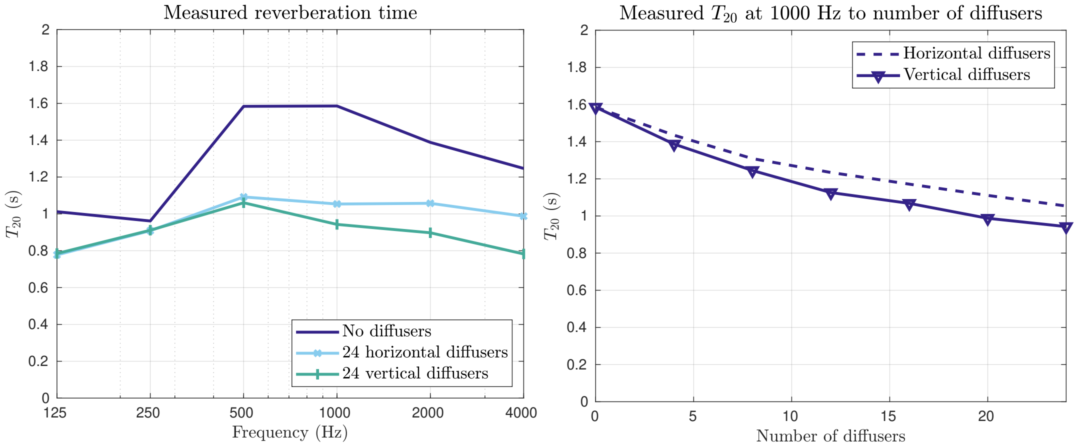
f (Hz)	**Floor**	**Doors**	**Scatterers**						**Walls**	**Absorptive Ceiling**
			Geo	**Uniform**	**Dir A, s_d**	**Dir B, s_d**	**Dir A, s_s**	**Dir B, s_s**		
125	0.05	0.1	0.1	0.12	0.12	0.1	0.12	0.1	0.13	0.13
250	0.05	0.1	0.1	0.15	0.15	0.1	0.15	0.1	0.15	0.15
500	0.05	0.1	0.15	0.23	0.23	0.35	0.23	0.5	0.05	0.05
1000	0.05	0.1	0.15	0.32	0.27	0.45	0.36	0.75	0.07	0.07
2000	0.05	0.15	0.25	0.37	0.31	0.45	0.42	0.8	0.05	0.05
4000	0.05	0.15	0.25	0.46	0.36	0.45	0.56	0.9	0.05	0.05

4. Results

4.1. Measurements

The measurements used in this study, as described in Section 3.1, are a subset of a larger set of measurements described in detail in [20–22,27]. Only the subset used in this study is shown here and only briefly discussed.

The plotted results, in Figure 9, indicate that the number of scatterers and their orientation had a significant impact on the reverberation time (T_{20}). In Figure 9a, the average measured reverberation time using no scatterers, and 24 horizontal and 24 vertical scatterers is shown. In Figure 9b, the reverberation time at 1000 Hz is shown for an increasing number of scatterers. It is clear that the introduction of scatterers led to a reduced reverberation time (which is consistent with previous research [7,29,30]) and that the reverberation time was further decreased when more scatterers were added.



(**a**) (**b**)

Figure 9. A brief overview of the measured reverberation time. In (**a**), the measured reverberation time is shown in the space with no scatterers, 24 vertically oriented scatterers and 24 horizontally oriented scatterers. In (**b**), the effects of increasing the number of scatterers is shown for 1000 Hz. In both graphs, it is evident that scatterers decrease the reverberation irrespective of orientation. However, vertically oriented scatterers reduce reverberation more at higher frequencies.

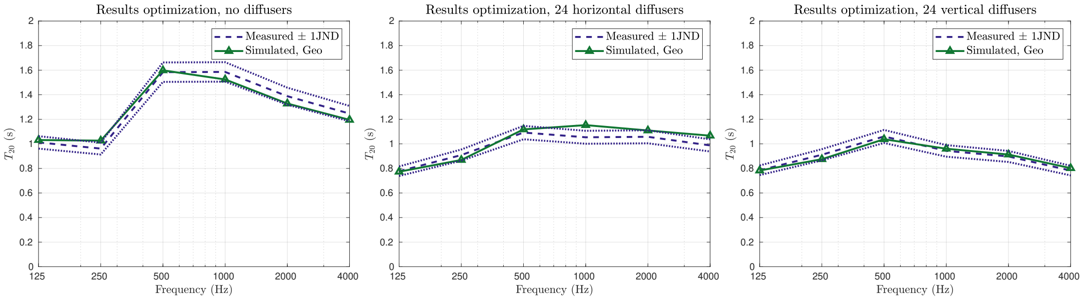
In this project, the orientation of scatterers is of particular interest. Based on the results in Figure 9a, the orientation had a measurable impact on the reverberation time at high frequencies but not at 500 Hz or below. Below 500 Hz, the wavelengths were large (about 70 cm and above), and the scatterers were not large enough to be expected to have a significant scattering effect on the sound field. The reverberation time reduction was instead due to the resonance absorption of the scatterers, which does not depend on their orientation. At 1000 Hz and above, there was very little absorption by the scatterers, but the acoustic energy was expected to be redirected by the scatterers to a high degree. The difference observed between horizontally and vertically oriented scatterers was because the vertically oriented scatterers redirected the acoustic energy towards the ceiling absorbers more efficiently.

4.2. Optimization Results

As mentioned in Section 3.3, an optimization method was used to obtain simulation parameters that produced simulation results in accordance with measurements. Although this is a good method to find parameter values that produce accurate simulation results, it is possible that the found parameter values are not accurate in relation to the actual surface properties. In this section, the results of the optimization procedure are briefly discussed, focusing on possible errors and how these may later affect the simulation results.

Some results of the optimization step are shown in Figure 10. As indicated in the graphs, the average simulated reverberation time is close to the measurements and mostly within 1 JND across all frequency bands, indicating that the determined parameters adequately simulate the average late decay in the space across all scatterer configurations.

At low frequencies, there are typically some resonance effects in rectangular spaces such as the one studied in this project. This can lead to spatial variations in the sound field that cannot be modelled using ray tracers, as they typically do not model phase effects. In this case, the least squares cost function used for optimization is likely to push the simulation results in each position towards the measured spatial average. Although this may be appropriate due to the limitations of ray tracers, it may also lead to a more homogeneous sound field in the simulation. In terms of parameter values, it is then likely that the estimated absorption parameters vary less in space and scattering parameters are overestimated. If such errors are made, it is likely that the simulations underestimate the effects of the orientation of the scatterers. Since this issue occurs primarily at low frequencies, where no effects could be measured due to the orientation of the scatterers, it is not expected to lead to any major problems.



(**a**) (**b**) (**c**)

Figure 10. Some results from the optimization step, for no scatters (**a**), 24 horizontal scatterers (**b**) and 24 vertical scatterers (**c**). It is shown that the optimized simulations (green) are within about 1 JND of the measurements in most cases.

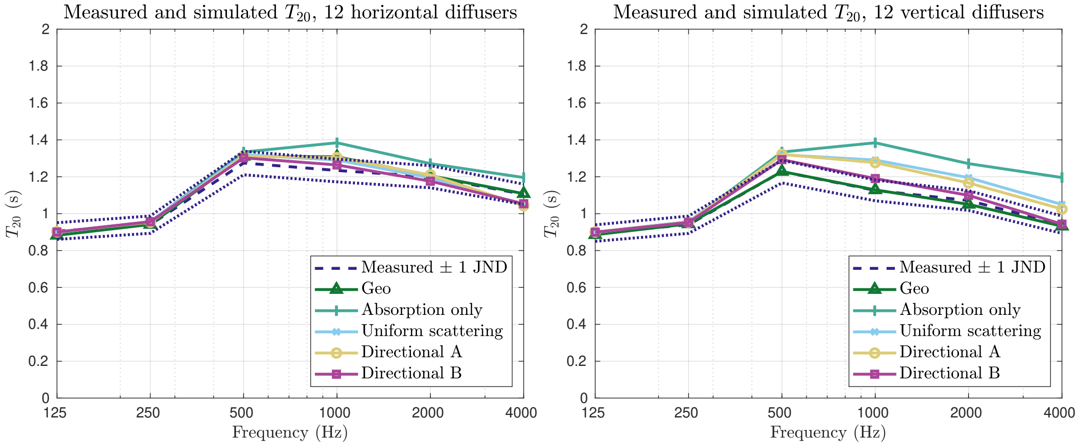
Revising the estimated absorption parameters (see Table 1) shows that the ceiling absorption parameter is much smaller than what is typically reported as table values for this type of ceiling. This is not necessarily an optimization error. Table values for absorption parameters are typically measured at random incidence. In a rectangular space with an absorptive ceiling, acoustic energy does not follow a random-incidence distribution towards the ceiling over time. Over time, mostly sound energy flowing in a direction parallel or approximately parallel to the ceiling remains [31]. Due to the direction of propagation of this acoustic energy, it impacts the ceiling at a large angle and is absorbed at a lesser degree than what is measured at random incidence. This leads to a practical absorption coefficient that is significantly smaller than table values. When scatterers are included, the apparent absorption of the ceiling is higher. It is, therefore, important that such cases are included in the optimization step.

4.3. Simulation Results

The measurements were compared to the simulation results from the five scattering models described in Section 3.2. In Figure 11, the averaged reverberation time for simulations and measurements using 12 horizontally (Figure 11a) and 12 vertically (Figure 11b) oriented scatterers are shown. As expected, the different simulation setups performed differently. In particular, the Geo model matches the simulation results better than the others. This was expected, since it was used for parameter optimization.

At the frequencies of 125 Hz and 250 Hz, there were no significant differences among the simulation methods for either of the configurations. This supports the earlier conclusion that absorption is dominant at these frequencies, particularly since the Absorption only strategy performed on par with the other models. At higher frequencies, there were some differences among the simulation strategies. The Absorption only model overestimated the reverberation time for both vertical and horizontal scatterers, due to the lack of scattering.

In Table 3, estimates of the average error in T_{20}, EDT and C_{50} are shown for each of the simulation procedures across all positions, frequency bands and scatterer configurations. The error is estimated as the square root of the average squared error, for all positions, scatterer configurations and frequency bands. The differences between the different simulation strategies are small in EDT and C_{50} and should not be used to make definitive statements.



(**a**) (**b**)

Figure 11. Graphs showing the measured and simulated reverberation time in the space with 12 horizontally (**a**) and vertically (**b**) oriented scatterers, respectively. In (**b**), the measured reverberation time is shorter as the scatterers, are vertically oriented. Not all simulated data sets are able to replicate this behaviour.

Table 3. Estimated error between simulations and measurements for different simulation methods and configurations. The variations in error for early decay time and speech clarity are small.

Average Error, Measurements vs. Simulations	**T_{20} (s)**	**EDT (s)**	**C_{50} (dB)**
Geometric	0.074	0.178	2.025
Absorption only	0.153	0.175	2.076
Uniform scattering	0.100	0.178	2.069
Directional scattering A	0.097	0.175	2.058
Directional scattering B	0.078	0.178	2.049

In general, the Geo simulation strategy worked best. This was to be expected, as it was used for parameter optimization. For reverberation time and C_{50}, the directional scattering models are better than the uniform models. In particular, Directional scattering B with the adjusted scattering parameters performed well in reverberation time. While the parameter adjustments are expected to lead to a better adherence between simulation and measurements, the fact that the directional scattering algorithm can be used to achieve significant improvement compared with the uniform scattering algorithm is very promising. It may also indicate that larger scattering parameters should be used to achieve good modelling performance when the directional scattering algorithm is used.

EDT does not follow the same pattern as the reverberation time or speech clarity. In this case, the models with the lowest practical scattering seemed to work best, namely, the Absorption only model and the Directional scattering A model. EDT is closely related to early reflections, and in real spaces, early reflections can be quite well defined and can often be well modelled using image source methods, without any scattering at all. The results found in this work may indicate that early reflections and the very early parts of the impulse response are better modelled without scattering. However, the variations depending on the method are very small, and no significant conclusions should be drawn from this result.

The results presented graphically were all spatially averaged, which is useful for giving a good overview of the overall tendencies in the results. In this project, the spatial variations were found to be independent of the modelling strategy used for the scatterers, and these results are consequently not explicitly presented. They are, however, included in the error estimates presented in Table 3.

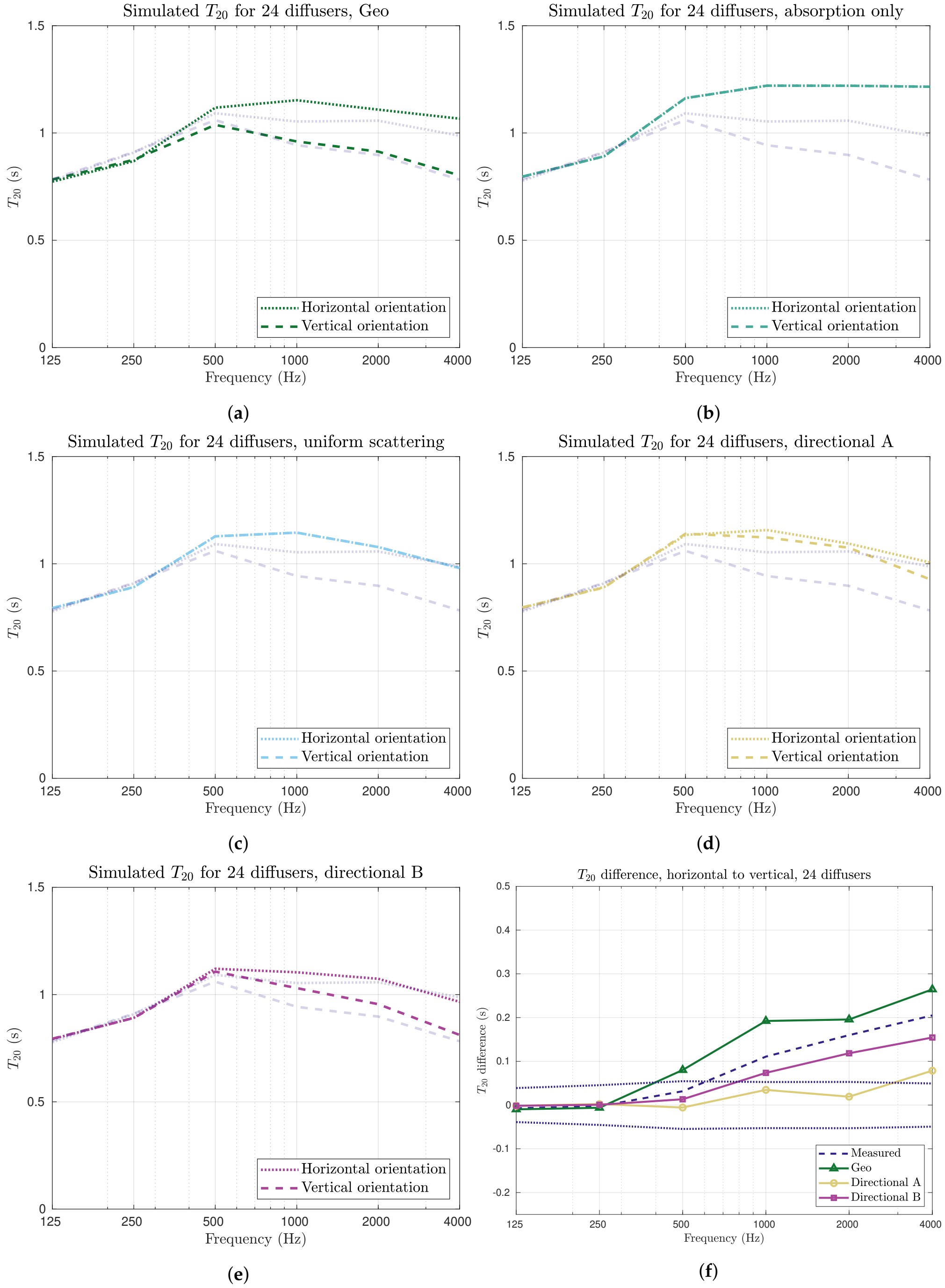
Finally, the differences in reverberation time caused by the orientation of the scatterers are examined in more detail. In Figure 12, the reverberation time with 24 horizontally or vertically oriented scatterers is shown for each simulation method and compared to measurements (shaded). In these results, it is clear that the differences that can be noted in measurements are not present in the non-directional algorithms. The Absorption only model clearly overestimated the reverberation time compared with measurements due to its lack of scattering. The Uniform scattering algorithm estimated the results with horizontally oriented scatterers fairly well but had too little scattering in the case with vertically oriented scatterers.

For the three configurations that can emulate the directionality of the scatterers, the effects of scatterer orientation could be seen in all cases. The Geo configuration performed well but slightly overestimated the reverberation time for the horizontally oriented scatterers. It is possible that this occurred due to edge effects from the mounting of the scatterers affecting the measurements. When the scatterers are mounted to redirect acoustic energy horizontally, the top and bottom edges (as seen in Figure 6b) are visible and may cause increased scattering in the vertical direction. When they are instead mounted vertically, those edges are mostly covered by other scatterers and do not contribute to scattering. This could lead to a difference in which scattering parameters are suitable for the scatterers depending on their orientation. Such an adjustment was not performed in this case, and it is possible that the scattering parameter was underestimated for horizontally oriented scatterers, leading to an overestimation of the reverberation time.

In the Directional scattering A strategy (Figure 12d), there was too little scattering overall. This is consistent with results presented earlier. It indicates that the method used for estimating the scattering parameters was not appropriate for the directional scattering algorithm. Directional scattering B (Figure 12e) showed similar problems but was overall much closer to the measured results. It is likely that further refinement or the use of an optimization strategy would further improve the adherence between simulation and measurements. However, this should be undertaken with care, as the quality of the scattering parameters is very difficult to evaluate without more research.

In Figure 12f, the differences between the vertical and horizontal scatterer orientation are presented and compared to the JND of measurements. It is expected that differences exceeding 1 JND correspond to an audible change in the sound environment of the space. The Geometric modelling method overestimated the differences between the vertical and horizontal orientations above 500 Hz. At higher frequencies, this might be explained by an underestimation of the scattering effects due to the horizontally oriented scatterers as discussed above. At 500 Hz, it is possible that this is explained by the resolution of the model leading to a larger impact of the scatterer orientation (see discussion in Section 1). However, such issues are not seen at lower frequencies, where the errors are expected to be more influential.

The scattering coefficients based on measurements did provide some difference between the two orientations, but it was comparatively small and not convincingly above 1 JND until 4000 Hz. Configuration B, where the scattering coefficients were adjusted, performed significantly better. In terms of algorithm performance, the differences that can be achieved using the directional algorithm are comparable to the differences seen using a geometric modelling procedure. This indicates that the algorithm can be useful as an alternative where geometric modelling is impractical or impossible due to exceedingly complex surfaces or issues in finding an appropriate spatial resolution.



(**a**) (**b**) (**c**) (**d**) (**e**) (**f**)

Figure 12. In (**a**–**e**), the simulated T_{20} results for 24 vertical and horizontal scatterers, respectively, are shown using the simulation setups described above. The measurements are shown as shaded graphs (see also Figure 9a). There is no difference between horizontal and vertical orientations for the Uniform and Absorption only setups. The simulated difference for Directional A is quite small. In (**f**), the T_{20} difference between horizontal and vertical configurations is shown, compared with 1 JND of measurements.

5. Discussion

In this study, a directional scattering algorithm for use in acoustic ray tracers is developed. The underlying model is presented, and the implementation is discussed. The algorithm is then used to simulate measurements where the effects of directional scattering are present. It was indicated that the developed algorithm was capable of replicating the measured differences for different orientations of 1D scatterers.

The scatterers used for measurements were found to have significant absorption at low frequencies. As such, the amount of energy reflected from the scatterers in this frequency range is small, and its distribution has a limited impact on the sound field. Since the effects of the distribution of reflected energy cannot be properly measured, it is not possible to evaluate the accuracy of the directional scattering algorithm in this case. Since it is theorized that the directional scattering algorithm may be preferable to geometric modelling of scatterers at low frequencies, further research using scatterers with less low-frequency absorption is needed.

As previously indicated, further investigation is needed to determine appropriate parameter ranges for the directional scattering algorithm. It is possible that experience with the scattering coefficient defined in the international standard could serve as an indication, but the results in this paper indicate that larger values may be appropriate for the directional scattering algorithm. The standard scattering coefficient can be interpreted as a measure of how much directional information is lost in reflection. In the directional scattering algorithm, this is not a valid interpretation of the scattering parameters. The larger scattering parameter, s_s, defines the amount of energy that is partially scattered, but some directional information is preserved for this energy. It is thus possible that it should be higher than the standard scattering coefficient to achieve the same amount of disorder after reflection. Similarly, the s_d parameter does not account for *all* lost information in the reflection. It is thus possible that it should be smaller than the standard scattering coefficient.

In the future, the reflection and scattering patterns as produced by the directional scattering algorithm should be revised further and compared to measurements or full numerical simulations. This is expected to offer insight into deficiencies and advantages of the model and be useful in terms of determining which scattering parameters may be appropriate.

6. Conclusions

The suggested algorithm for directional scattering in acoustic ray tracers is capable of replicating measured variations in RA parameters, as caused by changes in the orientation of 1D scatterers in a room with absorbent ceiling. This could be used for modelling such structures in the future. Further research is needed to evaluate how the method compares to geometric modelling at low frequencies and to determine suitable scattering parameters.

Author Contributions: Conceptualization, H.A. and E.N.; methodology, H.A.; software, H.A.; validation, H.A.; formal analysis, H.A.; investigation, H.A.; writing—original draft preparation, H.A.; writing—review and editing, E.N.; supervision, E.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Swedish Research Council, grant number 2016-01784. The APC was funded in part by Lund University.

Data Availability Statement: Restrictions apply to the availability of these data. Simulation results and code are available from the authors. Measurement data have been published elsewhere (see references) and may be requested from the authors of those publications.

Acknowledgments: Susanne Heyden for supervision and support Emma Arvidsson for measurements.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Savioja, L.; Svensson, P. Overview of geometrical room acoustic modeling techniques. *J. Acoust. Soc. Am.* **2015**, *138*, 708–730. [CrossRef] [PubMed]
2. Vorländer, M. *Auralization. Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 1–335. [CrossRef]
3. Cox, T.J.; D'Antonio, P. *Acoustic Absorbers and Diffusers*, 3rd ed.; Taylor & Francis Group: Boca Raton, FL, USA, 2017.
4. Kuttruff, H. *Room Acoustics*, 5th ed.; CRC Press: Boca Raton, FL, USA, 2009.
5. Cox, T.J.; Dalenback, B.I.; D'Antonio, P.; Embrechts, J.J.; Jeon, J.Y.; Mommertz, E.; Vorländer, M. A tutorial on scattering and diffusion coefficients for room acoustic surfaces. *Acta Acust. United Acust.* **2006**, *92*, 1–15.
6. Morse, P.M.; Ingard, K.U. *Theoretical Acoustics*; McGraw-Hill Book Company: New York, NY, USA, 1968.
7. Autio, H.; Vardaxis, N.G.; Bard Hagberg, D. The Influence of Different Scattering Algorithms on Room Acoustic Simulations in Rectangular Rooms. *Buildings* **2021**, *11*, 414. [CrossRef]
8. Dalenbäck, B.I. Modeling 1D-Diffusers—The Missing Link. 2011. Available online: http://www.catt.se/Lambert-1D-CATT.pdf (accessed on 18 October 2023).
9. Svensson, U.P.; Fred, R.I.; Vanderkooy, J. An analytic secondary source model of edge diffraction impulse responses. *J. Acoust. Soc. Am.* **1999**, *106*, 2331–2344. [CrossRef]
10. Lu, B.; Darmon, M.; Potel, C.; Zernov, V. Models Comparison for the scattering of an acoustic wave on immersed targets. *J. Phys. Conf. Ser.* **2012**, *353*, 012009. [CrossRef]
11. Lü, B.; Darmon, M.; Fradkin, L.; Potel, C. Numerical comparison of acoustic wedge models, with application to ultrasonic telemetry. *Ultrasonics* **2016**, *65*, 5–9. [CrossRef] [PubMed]
12. Chehade, S.; Kamta Djakou, A.; Darmon, M.; Lebeau, G. The spectral functions method for acoustic wave diffraction by a stress-free wedge: Theory and validation. *J. Comput. Phys.* **2019**, *377*, 200–218. [CrossRef]
13. Chen, L.; Zhang, Y.; Lian, H.; Atroshchenko, E.; Ding, C.; Bordas, S. Seamless integration of computer-aided geometric modeling and acoustic simulation: Isogeometric boundary element methods based on Catmull-Clark subdivision surfaces. *Adv. Eng. Softw.* **2020**, *149*, 102879. [CrossRef]
14. Cheng, S.; Wang, F.; Li, P.W.; Qu, W. Singular boundary method for 2D and 3D acoustic design sensitivity analysisImage 1. *Comput. Math. Appl.* **2022**, *119*, 371–386. [CrossRef]
15. Li, J.; Fu, Z.; Gu, Y.; Zhang, L. Rapid calculation of large-scale acoustic scattering from complex targets by a dual-level fast direct solver. *Comput. Math. Appl.* **2023**, *130*, 1–9. [CrossRef]
16. Shaaban, A.M.; Anitescu, C.; Atroshchenko, E.; Rabczuk, T. 3D isogeometric boundary element analysis and structural shape optimization for Helmholtz acoustic scattering problems. *Comput. Methods Appl. Mech. Eng.* **2021**, *384*, 113950. [CrossRef]
17. Wang, F.; Gu, Y.; Qu, W.; Zhang, C. Localized boundary knot method and its application to large-scale acoustic problems. *Comput. Methods Appl. Mech. Eng.* **2020**, *361*, 112729. [CrossRef]
18. Odeon A/S. ODEON Room Acoustics Software—User's Manual. 2020. Available online: https://odeon.dk/download/Version17/OdeonManual.pdf (accessed on 23 October 2023).
19. Dalenbäck, B.I. CATT-Acoustic v9.1 Powered by TUCT v2. 2022. Available online: https://www.catt.se/ (accessed on 18 October 2023).
20. Arvidsson, E.; Nilsson, E.; Hagberg, D.B.; Karlsson, O.J.I. The Effect on Room Acoustical Parameters Using a Combination of Absorbers and Diffusers—An Experimental Study in a Classroom. *Acoustics* **2020**, *2*, 505–523. [CrossRef]
21. Arvidsson, E.; Nilsson, E.; Bard-Hagberg, D.; Karlsson, O.J.I. Subjective Experience of Speech Depending on the Acoustic Treatment in an Ordinary Room. *Int. J. Environ. Res. Public Health* **2021**, *18*, 12274. [CrossRef] [PubMed]
22. Arvidsson, E.; Nilsson, E.; Bard Hagberg, D.; Karlsson, O.J.I. The Difference in Subjective Experience Related to Acoustic Treatments in an Ordinary Public Room: A Case Study. *Acoustics* **2021**, *3*, 442–461. [CrossRef]
23. Kuttruff, H. *Room Acoustics*, 6th ed.; CRC Press: Boca Raton, FL, USA, 2017.
24. *ISO 17497-1:2004*; Acoustics—Sound-Scattering Properties of Surfaces—Part 1: Measurement of the Random-Incidence Scattering Coefficient in a Reverberation Room. ISO: Geneva, Switzerland, 2004.
25. Givens, W. Computation of plain unitary rotations transforming a general matrix to triangular form. *J. Soc. Ind. Appl. Math.* **1958**, *6*, 26–50. [CrossRef]
26. Postma, B.N.J.; Katz, B.F.G. Creation and calibration method of acoustical models for historic virtual reality auralizations. *Virtual Real.* **2015**, *19*, 161–180. [CrossRef]
27. Arvidsson, E.; Nilsson, E.; Bard Hagberg, D.; Karlsson, O.J.I. Quantification of the Absorption and Scattering Effects of Diffusers in a Room with Absorbent Ceiling. *Buildings* **2021**, *11*, 612. [CrossRef]
28. *ISO 3382-1:2009*; Acoustics—Measurement of Room Acoustic Parameters—Part 1: Performance Spaces. ISO: Geneva, Switzerland, 2009.
29. Schröder, D. Physically Based Real-Time Auralization of Interactive Virtual Environments. Ph.D. Thesis, Institute of Technical Acousitcs, Aachen, Germany, 2011.

30. Zhu, X.; Kang, J.; Ma, H. The impact of surface scattering on reverberation time in differently shaped spaces. *Appl. Sci.* **2020**, *10*, 4880. [CrossRef]
31. Nilsson, E.; Arvidsson, E. An Energy Model for the Calculation of Room Acoustic Parameters in Rectangular Rooms with Absorbent Ceilings. *Appl. Sci.* **2021**, *11*, 6607. [CrossRef]

Paper D

acoustics

IMPACT FACTOR 2.1

CITESCORE 4.0

Article

An Iterative Ray Tracing Algorithm to Increase Simulation Speed While Maintaining Overall Precision

Hanna Autio, Nikolaos-Georgios Vardaxis and Delphine Bard Hagberg

Topic Collection
Featured Position and Review Papers in Acoustics Science

Edited by
Prof. Dr. Jian Kang

MDPI

https://doi.org/10.3390/acoustics5010019

acoustics

Article

An Iterative Ray Tracing Algorithm to Increase Simulation Speed While Maintaining Overall Precision

Hanna Autio *, **Nikolaos-Georgios Vardaxis** and **Delphine Bard Hagberg** †

Division of Engineering Acoustics, Lund University, P.O. Box 118, 221 00 Lund, Sweden
* Correspondence: hanna.autio@construction.lth.se
† Current address: Laboratory of Acousitcs, Soft Matter and Biophysics, Department of Physics and Astronomy, KU Leuven, Celestijnenlaan 200D, B3001 Heverlee, Belgium.

Abstract: Ray tracing is a frequently used method for acoustic simulations, valued for its calculation speed and ease of use. Although it is fast, there are no fully ray tracing-based real-time simulation methods or engines. Under real-time restrictions, ray tracing simulations lose precision and the variance inherent in the random simulation method has too much impact on the outcome. In this paper, an algorithm called iterative ray tracing is presented that reduces the negative effects of real-time restrictions by iteratively improving the initial calculation and increasing the precision over time. In addition, new estimates of the expected value and variance of ray tracing simulations are presented and used to show the iteration steps in the new algorithm reduce variance, while maintaining the expected value. Simulations using iterative ray tracing are compared to measurements and simulations using the classical ray tracing method, and it is shown that iterative ray tracing can be used to improve precision over time. Although more testing is needed, iterative ray tracing can be used to extend most ray tracing algorithms, in order to decrease the adverse effects of real-time restrictions.

Keywords: acoustic ray tracing; real-time acoustics; geometrical acoustics

Citation: Autio, H.; Vardaxis, N.-G.; Bard Hagberg, D. An Iterative Ray Tracing Algorithm to Increase Simulation Speed While Maintaining Overall Precision. *Acoustics* **2023**, *5*, 320–342. https://doi.org/10.3390/acoustics5010019

Academic Editor: Jian Kang

Received: 3 February 2023
Revised: 22 February 2023
Accepted: 10 March 2023
Published: 14 March 2023

1. Introduction

Ray tracing is one of the most common methods of acoustic simulations [1] and is used in popular commercial acoustic simulation software [2–4]. Among its advantages are its relatively low computational load and its close relationship to geometrical acoustic (GA) models for sound propagation [1,5,6]. The relatively low computational load means that ray tracing is a good candidate for acoustic simulations in virtual reality (VR) and other real-time applications. Additionally, acoustic ray tracing can be used to unlock the extreme computational capabilities of modern Graphics Processing Units (GPUs) for the purpose of acoustic simulations [7,8]. The ray tracer used in this study is implemented to run on the GPU.

Today, there exists a range of acoustic simulation software aimed at VR and acoustically immersive experiences. The most precise simulation tools achieve their accuracy by solving the wave equation for low frequencies, but cannot be considered fully real-time, as they rely on computing at least parts of the acoustic response in advance [9,10]. On the other hand, some tools aimed primarily at entertainment purposes, such as video games, can be very fast, but are found to be of questionable physical accuracy [11]. In between these extremes, there are at least two applications relying on geometrical acoustics that are capable of real-time simulation. One is developed at the Technical University of Aachen (RWTH Aachen), and called Room Acoustics for Virtual Environments (RAVEN) [12]. It implements image source, ray tracing, and radiosity techniques. A second example is EVERTims [13], which is an open-source solution that combines an image sources with a statistical model for the late reverberation. Image source modelling is computationally efficient if a low order of image sources is used, which explains their prevalence in real-time modelling techniques. In EVERTims, the order of image sources is increased over time, thus refining the simulation data. As of yet, there is no real-time acoustic simulation software relying only on ray tracing, despite its popularity as an element of room acoustic simulation software [2,3].

If an acoustic application is to be perceptually real-time, the simulated sound field must respond to changes in the virtual environment sufficiently fast. There is no firm consensus on what should be considered "sufficiently fast", but values in the range of 20–60 ms are called interactive in the literature [13,14]. The actual time limit depends on the actions of the listener and their attention, and a hard bound is, thus, hard to define. Within this time limit, the sound field must be updated to account for any changes in the environment. Depending on what these changes entail, the sound field update may consist only of impulse response filtering or the calculation of a completely new simulation [12]. The variations in time requirements indicate that there is a need for more adaptive simulation strategies.

Ray tracing is a prevalent room acoustic simulation tool, and the term defines both the model for sound propagation (geometrical acoustics) and the calculation method. The sound field is modelled as a set of plane waves, and typically only the energy content of the waves are considered [1,5]. With these assumptions, the sound field can be described as an infinite set of energy particles whose behavior is well-defined. This model for the sound field is called geometrical acoustics [6], and in the case of ray tracing, the acoustic field is estimated by Monte Carlo sampling of the energy particles (or rays) [15].

Monte Carlo methods are a class of simulation methods that rely on random sampling to calculate something that may be deterministic in nature. In the case of room acoustic simulations, the impulse response is defined by the wave equation and the space under consideration and is, thus, deterministic. It is estimated using random samples of energy particles and their paths. In practice, this means calculating the potential propagation paths for a large set of rays and determining their energy levels over time. As the rays intersect the position of the sound receiver, their energy is recorded and added to an estimate of the energetic impulse response. By sampling a large number of random ray paths, the deterministic impulse response can be estimated. Frequently, the impulse response is estimated in a given subset of positions, leading to a spatial discretization.

Although the impulse response is deterministic, the ray tracing simulation method itself is random. Consequently, the simulation outcome can be described by its statistical properties. The expected value of a simulation is the energetic impulse response defined by the geometrical acoustics model. The variance of the simulation outcome describes how much it may differ between calculations, and it can be used to gauge how accurate the calculation is. As is typical for Monte Carlo simulation methods, the expected value is independent of the number of samples whereas the variance decreases as the number of samples increase [16]. Since the variance should be minimized, it is often desirable to use as many rays as possible. However, the higher the number of rays, the higher the computational load and the longer the simulation time. Although the expected value and variance of ray tracing simulations typically cannot be calculated outright, accurate estimates provide insight into the simulation tool itself and can be used as a tool to evaluate simulation parameters, such as the number of rays used.

The lack of a real-time ray tracing option for acoustic simulations, despite its popularity, shows that there is a need for algorithms that allow for faster calculation. On the other hand, the simulation accuracy and precision should be maintained as far as possible, and it has been noted that the real-time limitations vary depending on the situation. A more flexible ray tracing simulation algorithm can be used to fill in both of these gaps.

In this paper, a ray tracing algorithm is presented, which allows for greater flexibility in finding the balance between simulation speed and precision. The algorithm, called *Iterative ray tracing* , can mitigate the issues caused by small sample sizes in ray tracing simulations when speed cannot be compromised with. It iteratively improves the simulation result until maximum precision is reached. The algorithm itself is easy to combine with other methods for improving the performance of ray tracers.

The algorithm is statistically equivalent to classical ray tracing. This is shown by developing a novel theoretical estimate of the expected value and variance of the energetic impulse response produced by ray tracing under diffused conditions. These estimates improve on earlier versions [14] by fully accounting for the randomness involved in ray tracing, and are compared to results from real simulations to ensure that they are accurate.

The following sections presents and explains iterative ray tracing, as well as formulates the novel estimates of the expected value and variance. From this foundation, the accuracy of the estimates are shown, and the functionality of the algorithm is demonstrated. This is followed by a brief discussion and conclusion.

2. Iterative Ray Tracing: Introduction

In real time acoustic simulation, the sound field presented to the listener should adapt to changes in the environment sufficiently fast that no delay can be perceived. When the listener, or user, is moving quickly through the simulated space or interacting with the environment in acoustically relevant ways (e.g., making noise, changing the environment by opening doors or similar), the real time requirement can be very restrictive. On the other hand, the user may, at times, refrain from interacting at all with the environment, instead remaining in one place and perceiving the surrounding space. In such situations, the time limit becomes almost irrelevant. Iterative ray tracing aims at improving the tools for finding a balance between these two extremes.

In the context of ray tracing, restrictive time limits impact the number of rays that can be used in the simulation. Reducing the number of rays does not change the expected outcome of the simulation, but it can increase variance to the point where the result is not usable. When the time limit imposes an upper bound on the number of rays that negatively affects the reliability of the simulation result, while not rendering it unusable, a possible solution is to iteratively improve the simulation result. In this way, the simulation can offer excellent precision, while still generating simulation results quickly enough to comply with real-time restriction [13]. This is the idea behind iterative ray tracing.

The calculated impulse response can be iteratively improved in cases when the simulated space has not changed since the last time frame. In other cases, the impulse response is no longer accurate and should be discarded. In the context of ray tracing, an intuitive way of improving the simulated impulse response is by increasing the number of rays. In iterative ray tracing, this is achieved by adding the results from a new ray tracing simulation to the existing, still accurate, impulse response. When the ray tracing simulations are configured in the right way, this is statistically equivalent to performing one single ray tracing simulation with a number of rays equal to the sums of the two separate simulations.

A flowchart illustrating the algorithm is shown in Figure 1. The ray tracing and auralization frameworks are presented as grey boxes, and their specific implementations are not considered. In this algorithm, the ray tracing engine is assumed to continuously produce impulse responses. New impulse responses (grey oval) are passed to the iterative algorithm, which is responsible for determining what to do next (yellow rhombus). If the acoustic field has changed significantly since the last frame, the newly produced impulse response is passed directly to auralization. If it has not, it is combined with the existing impulse response which may, in turn, consist of the result of multiple previous calculations, and then it is passed to the auralization engine. The combination of impulse responses is a central part of the iterative ray tracing algorithm. The conditions under which it can be performed and how it should be implemented are described in Section 3.2.

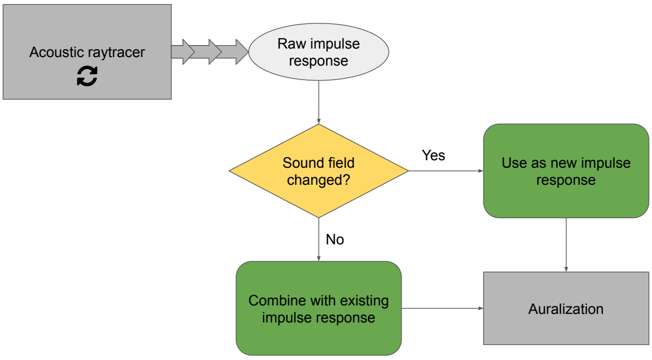


Figure 1. Flowchart illustrating the suggested algorithm. The ray tracer runs continuously, producing new simulated impulse responses (grey oval) as often as possible. For each new impulse response, the iterative algorithm determines whether the sound field has changed significantly or not since the last impulse response (yellow rhombus). If it has not, the simulation result is used to iteratively improve the existing impulse response. If it has, the simulation result is taken as a new impulse response. The final chosen impulse response is then passed on to the auralization engine.

The iterative algorithm is responsive to the events in the environment, in that speed is prioritized when needed and higher quality calculations are provided when possible. In general, when a fully new calculation is performed, it coincides with some significant event in the environment, which, in turn, may distract the listener from possible momentary acoustic variations, due to sample size [17]. On the other hand, when a listener is stationary and inactive, the sound simulation is reliable and realistic.

When implementing this algorithm, a central aspect is determining the criteria for when a new simulation is needed. When particular events occur, such as a door opening or closing, it can be straightforward to determine whether it has occurred and the simulation should be updated. In many cases, such as the user moving to a new location, the situation is less clear. Typically, a human standing still is not entirely still, but makes very small movements, due to both breathing and heart beats. Such movements should not lead to the previous simulation being discarded, as they do not cause audible changes to the sound field. On the other hand, if the simulation is only updated when audible changes occur, the updates are naturally audible, rather than imperceptible.

In practice, this leads to a spatial discretization. Its resolution depends on how acoustically homogenous the simulated space is. The Nyquist–Shannon sampling theorem suggests a spatial resolution of about 4 cm, given a maximum frequency of 4 kHz, but research findings suggest that the human hearing's spatial resolution for room acoustic variations is less fine [18,19]. Accurate estimates of the appropriate spatial resolution are yet to be determined, but are likely to be on the order of 20–75 cm. The spatial discretization can be based on the distance to the most recently calculated impulse response.

In the following section, it is shown that the iterative ray tracing algorithm can be implemented such that it has the same expected value and variance as the classical algorithm.

3. Expected Value and Variance of Ray Tracing Simulations

Iterative ray tracing relies on the fact that separate ray tracing simulations can be combined to reduce variance while maintaining the expected value. By studying the expected value and variance of impulse responses generated by ray tracing, as well as their sums, this prerequisite can be evaluated. While it is not possible to derive a general closed-form expression, estimates can be formed by making some strong assumptions on the distribution of sound energy in the simulated space. In this section, such estimates are formed for both classical and iterative ray tracing. The discussion in this section is applicable to ray tracers in general and is specific to the iterative algorithm only when specified.

While the estimates of the expected value and variance of ray tracing simulations exist [14], they do not take the full stochastic nature of Monte Carlo simulations into account. This leads to a significant simplification of the expressions and associated calculations, but may also have a significant impact on the accuracy of the estimates. The estimates derived in this paper do not rely on such assumptions and are, consequently, more formally accurate. They may, therefore, lead to an increased understanding of the behavior of ray tracing simulations and may be used to evaluate the accuracy of the earlier, simplified expressions. The derived expressions are compared both to earlier estimates and to real simulation outcomes, to evaluate their accuracy.

Theoretical estimates of the expected value and variance of a ray tracing simulations are a useful tool in learning more about ray tracing simulations, but can, in general, not be used as a tool to study the real sound field. In particular, the estimated variance only relates to the simulation outcome and not the real sound field in any way. Since ray tracing is a Monte Carlo simulation method, the expected result of a ray tracing simulation is the impulse response predicted by the GA model. However, the theoretical estimate of the expected value of the impulse response relies on strict assumptions on the sound field, which are generally never fulfilled in real spaces. Consequently, the expression for the expected value should not be used as a substitute for a real ray tracing simulation. In this paper, it is used to compare the expected value of different simulation configurations.

The most central assumption in the derivation of the estimates below is the assumption of diffuseness. In this context, a diffused sound field is a sound field which is described entirely by its statistical properties. Real sound fields can be more or less diffused, where a more diffused sound field has a more even distribution of sound energy with little to no net flow of acoustic energy. In less diffused sound fields, there is at least some net flow of acoustic energy caused by properties of the surrounding space itself. When making general estimates of the sound energy, no information is known about the space and the assumption of diffuseness becomes a best-guess of the distribution of sound energy. Since the expressions derived in this paper should be generally applicable, it is assumed that the sound field is sufficiently diffused to be described by its statistical properties.

Under some circumstances, the assumption of diffuseness is not fulfilled. In particular, in the very early parts of the impulse response, the sound field is not diffused. Immediately when a sound impulse is emitted, there is a concentration of sound energy at that position, and the sound field cannot be described as diffused before some amount of diffusion has occurred. Other factors that affect the level of diffuseness are the simulated space and the amount of surface scattering. In general, very geometrically regular spaces, low surface scattering, or very unevenly distributed absorption can lead to a sound field that is highly non-diffused. In these cases, the estimates derived in this section are less accurate.

Finally, some comments are made regarding frequency dependence and scattering and absorption parameters. In general, the behavior of low and high frequency acoustical waves differ significantly and this is typically implemented in ray tracers by frequency-dependent surface absorption and scattering. In the derivation of the expected value and variance, however, it is assumed that the sound field is diffused for all frequencies. A consequence of this is that the scattering parameter, which would normally relate to the amount of diffuseness in the sound field; however, is not considered in the calculations below. In terms of the absorption coefficient, the assumption of diffuseness ensures that an average absorption coefficient for the walls of the enclosure is adequate. It is, consequently, assumed to be constant for all surfaces in the sequel. In terms of frequency dependence, the theoretical estimates derived in this paper, thus, only takes it into account in the definition of surface and air absorption coefficients.

3.1. Classical Ray Tracing Simulation: Expected Value and Variance

In this section, an estimate is derived for the expected value and variance of the energetic impulse response found by ray tracing simulation. The energetic impulse is discretized in the time domain and contains the energy detected at the listener position for each time step. The expected value and variance is, consequently, also defined for each time frame. It is assumed that the ray tracing simulations take place in a closed volume, where acoustic energy is lost only by surface or air absorption. It is further assumed that surface absorption is constant and non-zero, to ensure that the acoustic energy decays over time.

We find that in time step k, the energy detected E_k can be described as

$$E_k = \sum_{j=1}^{N} i_{j,k} e_{j,k}, \tag{1}$$

where $i_{j,k}$ is the number of times ray j intersects the detector in time interval k, and $e_{j,k}$ is the energy of ray j in time interval k. N is the total number of rays in the simulation. Note that the energy of each ray is assumed to be constant during the full time step. This assumption is acceptable if the time step Δt is sufficiently small.

If all rays are generated the same way, with constant energy and a randomly generated direction, they have identical distribution, and the expected values $\mathbb{E}\left[i_{j,k}\right]$ and $\mathbb{E}\left[e_{j,k}\right]$ are constant for all values j. Furthermore, it is assumed that $i_{j,k}$ and $e_{j,k}$ are mutually independent. This is true if the energy of each ray is unrelated to how likely it is to be found close to the detector. For example, this assumption would not necessarily hold if the detector is placed close to several absorptive surfaces, as rays close to absorbers are more likely to have lower energy. Since we consider the sound field to be diffuse; however, we can assume that $i_{j,k}$ and $e_{j,k}$ are independent. We can then write the expected value $\mathbb{E}[E_k]$

$$\mathbb{E}[E_k] = \mathbb{E}\left[\sum_{j=1}^{N} i_{j,k} e_{j,k}\right] = \sum_{j=1}^{N} \mathbb{E}\left[i_{j,k}\right] \mathbb{E}\left[e_{j,k}\right] = N \mathbb{E}\left[i_{j,k}\right] \mathbb{E}\left[e_{j,k}\right]. \tag{2}$$

The number of times a given ray intersects the detector in a given time interval, $i_{j,k}$, is Poisson distributed. In fact, $i_{j,k} \in \mathrm{Po}\left(\bar{n}\frac{S_d}{S}\Delta t\right)$, where S_d is the surface area of the detector (given by $4\pi r_d^2$ if it is modelled as a sphere with radius r_d), S is the total surface area of the modelled space, and $\bar{n}$ is the average rate of reflection for a ray. In a diffused field, $\bar{n} = \frac{cS}{4V}$, where V is given by the volume of the space and c is the speed of sound [6]. This yields, for expected value $\mathbb{E}$ and variance $\mathbb{V}$,

$$\mathbb{E}\left[i_{j,k}\right] = \mathbb{V}\left[i_{j,k}\right] = \bar{n}\frac{S_d}{S}\Delta t = \bar{i}, \tag{3}$$

where the detection rate $\bar{i}$ has been introduced to simplify the expressions. It is the expected number of times a ray hits the detector in a single time step. Note that $\bar{i}$ is independent of j, as has already been discussed, but also independent of time k. In most ray tracing simulations, rays are eventually terminated and no longer detectable once their energy falls below some threshold. The detection rate should, thus, tend to zero over time. In the later parts of the impulse response, assuming a constant detection rate $\bar{i}$ may lead to an inflated estimate of the detected energy, as compared to the results of the simulation. If the energy threshold for termination is set sufficiently low in the simulation, this does not have a significant impact on the accuracy of the estimate.

The energy of each ray, $e_{j,k}$, decreases over time as energy is absorbed from surface reflections and air absorption. The latter is assumed to be deterministic, whereas energy absorbed from surface reflections depend on the number of reflections a given ray has undergone. It is found that

$$e_{j,k} = e_0 \mathrm{e}^{-\gamma c k \Delta t}(1-\alpha)^{m_{j,k}} = e_0 a_k (1-\alpha)^{m_{j,k}}, \tag{4}$$

where e_0 is the initial energy of an emitted ray, α is the absorption parameter of the walls of the enclosure (here assumed to be constant), $\mathrm{e}^{-\gamma c k \Delta t} = a_k$ accounts for air absorption, and $m_{j,k}$ is the number of surface reflections ray j has undergone at time step k. Since the number of reflections depends on the random ray path, it is stochastic and the energy is also stochastic. The number of reflections a ray has undergone at a given time is a Poisson process and $m_{j,k} \in \mathrm{Po}(\bar{n} k \Delta t)$. Accordingly, $\mathbb{E}\left[m_{j,k}\right] = \mathbb{V}\left[m_{j,k}\right] = \bar{n} k \Delta t = \bar{m}_k$, where the expected number of reflections a given ray has undergone at time k, $\bar{m}_k$, has been introduced. This number increases as k increases, but is, again, independent of j.

The ray energy $e_{j,k}$ is stochastic, as it depends on the path of the given ray, which is randomly generated. However, in a completely diffused sound field, the rays would be evenly distributed in space and have equal energy. It is, thus, possible to assume that the energy decay of each ray is deterministic and determined by models for diffused sound decay in enclosures, so that $m_{j,k} \triangleq \bar{m}_k$ and $\mathbb{V}\left[e_{j,k}\right] = 0$. With this assumption, the mathematical operations become significantly more simple and compact, and it has been used in previous derivations [14]. In this paper, the stochastic nature of the energy levels is, instead, dealt with by introducing a Taylor expansion. Either of these solutions are approximations that introduce some level of uncertainty.

As mentioned, the ray energy is a function of a stochastic variable in our model, $e_{j,k} = e_{j,k}\left(m_{j,k}\right)$. An estimate of $\mathbb{E}\left[e_{j,k}\right]$ can be formed by using a Taylor expansion about the expected value of $m_{j,k}$. The general expression for the second order Taylor expansion of the expected value of a function of a stochastic variable $\mathbb{E}[f(x)]$ about the expected value of $x = \mu$ is:

$$\mathbb{E}[f(x)] \approx \mathbb{E}\left[f(\mu) + f'(\mu)(x-\mu) + \frac{1}{2}f''(\mu)(x-\mu)^2\right] = f(\mu) + \frac{1}{2}f''(\mu)\mathbb{V}[x], \tag{5}$$

using the fact that $\mathbb{E}[x-\mu] = 0$ and $\mathbb{V}[x] = \mathbb{E}\left[(x-\mu)^2\right]$.

Inserting $e_{j,k} = e_{j,k}\left(m_{j,k}\right)$ for f and using $\mathbb{E}\left[m_{j,k}\right] = \mathbb{V}\left[m_{j,k}\right] = \bar{m}_k$, we obtain

$$\mathbb{E}\left[e_{j,k}\right] \approx e_0 a_k (1-\alpha)^{\bar{m}_k} + \frac{1}{2} e_0 a_k (1-\alpha)^{\bar{m}_k} \ln^2(1-\alpha)\bar{m}_k. \tag{6}$$

We define the dimensionless constant $\hat{\alpha} = \ln(1-\alpha)$ and insert Equations (3) and (6) into Equation (2), obtaining

$$\mathbb{E}[E_k] = N\mathbb{E}\left[i_{j,k}\right]\mathbb{E}\left[e_{j,k}\right] \approx N e_0 a_k \bar{i}\left(\mathrm{e}^{\hat{\alpha}\bar{m}_k} + \frac{\mathrm{e}^{\hat{\alpha}\bar{m}_k}}{2}\hat{\alpha}^2 \bar{m}_k\right). \tag{7}$$

This is an estimate of the expected energy in the energetic impulse response produced by a ray tracer. If $Ne_0 = 1 \Leftrightarrow e_0 = \frac{1}{N}$, the total energy emitted in the impulse is 1, and the impulse response is normalized. For this case, we set $\mathbb{E}[E_k] = \mathbf{E}_k$. When the initial ray energy is determined this way, the expected value is independent of the number of rays; additionally, the impulse response neither amplifies nor diminishes sounds emitted in the space.

The variance of the energetic impulse response can be estimated using a Taylor expansion in a similar way. Firstly, we note that its variance can be calculated by

$$\mathbb{V}[E_k] = \mathbb{V}\left[\sum_{j=1}^{N} i_{j,k} e_{j,k}\right] = N\left(\mathbb{V}\left[i_{j,k}\right]\mathbb{V}\left[e_{j,k}\right] + \mathbb{V}\left[i_{j,k}\right]\mathbb{E}^2\left[e_{j,k}\right] + \mathbb{V}\left[e_{j,k}\right]\mathbb{E}^2\left[i_{j,k}\right]\right), \tag{8}$$

if we assume that $\mathbb{COV}[E_k, E_l]$ for all $k \neq l$, or that they are mutually independent. Since the distribution of $i_{j,k}$ is known from Equation (3), we find that $\mathbb{V}\left[i_{j,k}\right] = \bar{i}$, $\mathbb{E}^2\left[i_{j,k}\right] = \bar{i}^2$

Using our earlier approximation of $\mathbb{E}\left[e_{j,k}\right]$, we determine

$$\mathbb{E}^2\left[e_{j,k}\right] \approx e_0^2 a_k^2 \mathrm{e}^{\hat{\alpha}2\bar{m}_k}\left(1 + \hat{\alpha}^2\bar{m}_k + \frac{1}{4}\hat{\alpha}^4\bar{m}_k^2\right). \tag{9}$$

In addition, an estimate of $\mathbb{E}\left[e_{j,k}^2\right]$ needs to be constructed. This is achieved by Taylor expansion,

$$\begin{aligned}\mathbb{E}\left[f^2(x)\right] &\approx \mathbb{E}\left[f^2(\mu) + 2f(\mu)f'(\mu)(x-\mu) + 2\frac{(x-\mu)^2}{2}\left(f'^2(\mu) + f(\mu)f''(\mu)\right)\right] \\ &= f^2(\mu) + \left(f'^2(\mu) + f(\mu)f''(\mu)\right)\mathbb{V}[x].\end{aligned} \tag{10}$$

Using $e_{j,k} = e_{j,k}(m_{j,k})$, $\mathbb{E}\left[m_{j,k}\right] = \mathbb{V}\left[m_{j,k}\right] = \bar{m}_k$ and Equation (9), an estimate of the variance $\mathbb{V}\left[e_{j,k}\right]$ is found as

$$\mathbb{V}\left[e_{j,k}\right] \approx e_0^2 a_k^2 \mathrm{e}^{\hat{\alpha}2\bar{m}_k}\left(\bar{m}_k\hat{\alpha}^2 - \frac{\hat{\alpha}^4}{4}\bar{m}_k^2\right) \tag{11}$$

Finally, Equations (9) and (11) are inserted into Equation (8) to arrive at an estimate for $\mathbb{V}[E_k]$:

$$\begin{aligned}\mathbb{V}[E_k] &= N\left(\mathbb{V}\left[i_{j,k}\right]\mathbb{V}\left[e_{j,k}\right] + \mathbb{V}\left[i_{j,k}\right]\mathbb{E}^2\left[e_{j,k}\right] + \mathbb{V}\left[e_{j,k}\right]\mathbb{E}^2\left[i_{j,k}\right]\right) \\ &\approx N e_0^2 a_k^2 \mathrm{e}^{\hat{\alpha}2\bar{m}_k}\bar{i}\left(1 + 2\bar{m}_k\hat{\alpha}^2 + \bar{i}\bar{m}_k\hat{\alpha}^2 - \frac{\hat{\alpha}^4}{4}\bar{m}_k^2\bar{i}\right)\end{aligned} \tag{12}$$

This is an estimate of the variance of the energetic impulse response produced by ray tracing simulations. Reviewing the expression, it seems that the total variance increases as the number of rays N increases. If we, instead, consider the normalized impulse response $\mathbf{E}_k$, the case when the product $Ne_0 = 1 \Leftrightarrow e_0 = \frac{1}{N}$, we find that

$$\mathbb{V}[\mathbf{E}_k] = \frac{1}{N}a_k^2 \mathrm{e}^{\hat{\alpha}2\bar{m}_k}\bar{i}\left(1 + 2\bar{m}_k\hat{\alpha}^2 + \bar{i}\bar{m}_k\hat{\alpha}^2 - \frac{\hat{\alpha}^4}{4}\bar{m}_k^2\bar{i}\right) = \frac{1}{N}V_k, \tag{13}$$

where V_k has been introduced for the factor of the variance that is independent of the number of rays. The total variance, thus, decreases as the number of rays increase.

In conclusion, an estimate for the expected value and variance of the outcome of acoustic ray tracing simulations has been derived, as shown in Equations (7) and (12). When the initial ray energy e_0 is set, so that $Ne_0 = 1$, the expected value is independent of and the variance is inversely proportional to the number of rays.

In order to show that the iterative ray tracing algorithm can be used to further decrease variance, while maintaining the expected value, it is reviewed in the following section.

3.2. Iterative Ray Tracing: Expected Value and Variance

In this section, the estimates of the expected value and variance of the iterative ray tracing algorithm are derived. From those estimates, the method for combining the impulse responses from distinct ray tracing simulations is more closely defined to ensure that the iterative ray tracing algorithm has the desired expected value and variance. It is shown that, using the method presented in this paper, the iterative ray tracing algorithm has the same expected value as the classical algorithm. Initially, only the combination of two simulations are considered. The results are then generalized to multiple simulations.

Consider the sum of two separate simulations, each consistent with the expression for the energetic impulse response in Equation (1). In general, the variables related to the secondary simulations are denoted as $\cdot^*$, and variables from the iterative algorithm are marked with an I. The expected value for the detected energy in each time step using the iterative algorithm, E_k^I, then becomes

$$E_k^I = \sum_{j=1}^{N} i_{j,k} e_{j,k} + \sum_{l=1}^{N^*} i_{l,k}^* e_{l,k}^*. \tag{14}$$

For the expected value and variance, it is known that

$$\begin{aligned} \mathbb{E}[x+y] &= \mathbb{E}[x] + \mathbb{E}[y], \\ \mathbb{V}[x+y] &= \mathbb{V}[x] + \mathbb{V}[y] + 2\mathbb{COV}[x,y], \end{aligned} \tag{15}$$

where $\mathbb{COV}$ denotes the covariance. Since the initial ray tracing step is generated randomly in the same way as the initial simulation, the expressions derived in Equations (7) and (12) can be used to determine the expected value and variance of the secondary sum in Equation (14). Regarding the covariance, it describes how the random variations in the two simulations are connected. When the two simulations are *randomly* and *independently* initialized, the covariance term can generally be discarded. It is, consequently, important to ensure that, for example, the same random seed or ray generation pattern is not used for both simulations.

We find that, using Equations (14), (7), and (12)

$$\begin{aligned} \mathbb{E}\left[E_k^I\right] \approx & N e_0 a_k \bar{i} \left(\mathrm{e}^{\hat{\alpha}\bar{m}_k} + \frac{\mathrm{e}^{\hat{\alpha}\bar{m}_k}}{2} \hat{\alpha}^2 \bar{m}_k \right) \\ & + N^* e_0^* a_k \bar{i}^* \left(\mathrm{e}^{\hat{\alpha}\bar{m}_k^*} + \frac{\mathrm{e}^{\hat{\alpha}\bar{m}_k^*}}{2} \hat{\alpha}^2 \bar{m}_k^* \right), \end{aligned} \tag{16}$$

$$\begin{aligned} \mathbb{V}\left[E_k^I\right] \approx & N e_0^2 a_k^2 \mathrm{e}^{\hat{\alpha} 2\bar{m}_k} \bar{i} \left(1 + 2\bar{m}_k \hat{\alpha}^2 + \bar{i}\bar{m}_k \hat{\alpha}^2 - \frac{\hat{\alpha}^4}{4} \bar{m}_k^2 \bar{i} \right) \\ & + N^* e_0^{*2} a_k^2 \mathrm{e}^{\hat{\alpha} 2\bar{m}_k^*} \bar{i} \left(1 + 2\bar{m}_k^* \hat{\alpha}^2 + \bar{i}^* \bar{m}_k^* \hat{\alpha}^2 - \frac{\hat{\alpha}^4}{4} \bar{m}_k^{*2} \bar{i}^* \right). \end{aligned} \tag{17}$$

If the simulated space has not changed significantly between the two simulations, it can be assumed that $\bar{i} \approx \bar{i}^*$ and $\bar{m}_k \approx \bar{m}_k^*$. This mean that the expected number of times a given ray hits the detector $\bar{i}$ and the expected number of reflections a ray has undergone at time k, $\bar{m}_k$ is the same in the two simulations. Using this assumption, we can express Equations (16) and (17) as

$$\mathbb{E}\left[E_k^I\right] \approx (N e_0 + N^* e_0^*) a_k \bar{i} \left(\mathrm{e}^{\hat{\alpha}\bar{m}_k} + \frac{\mathrm{e}^{\hat{\alpha}\bar{m}_k}}{2} \hat{\alpha}^2 \bar{m}_k \right) = (N e_0 + N^* e_0^*) \mathbf{E}_k, \tag{18}$$

$$\mathbb{V}\left[E_k^I\right] \approx \left(N e_0^2 + N^* e_0^{*2} \right) a_k^2 \mathrm{e}^{\hat{\alpha} 2\bar{m}_k} \bar{i} \left(1 + 2\bar{m}_k \hat{\alpha}^2 + \bar{i}\bar{m}_k \hat{\alpha}^2 - \frac{\hat{\alpha}^4}{4} \bar{m}_k^2 \bar{i} \right) = \left(N e_0^2 + N^* e_0^{*2} \right) V_k, \tag{19}$$

where $\mathbf{E}_k$, the expected value of the normalized impulse response and V_k, and the portion of the variance that is independent of the number of rays (see Equation (13)) has been reintroduced.

The results presented above can easily be generalized to the case where more than two simulations are used, and we find that for the νth iterative step,

$$\mathbb{E}[E_{k,\nu}] = \left(\sum_{\kappa=1}^{\nu} N_\kappa e_{0,\kappa} \right) \mathbf{E}_k, \tag{20}$$

$$\mathbb{V}[E_{k,\nu}] = \left(\sum_{\kappa=1}^{\nu} N_\kappa e_{0,\kappa}^2 \right) V_k. \tag{21}$$

In the case of a single simulation, the initial ray energy is typically set so that $Ne_0 = 1$, as this produces an impulse response that is normal, in the sense that it does not artificially amplify or diminish sound energy. When this is the case, $\mathbb{E}[E_k] = \mathbf{E}_k$. Since the iterative ray tracing algorithm should not change the expected outcome of the simulation, the resulting summed impulse response $E_{k,\nu}$ in Equation (20) must be normalized by $\sum N_\kappa e_{0,\kappa}$ or the total emitted energy in all simulations.

We find that

$$\mathbb{E}\left[\frac{1}{\sum_{\kappa=1}^{\nu} N_\kappa e_{0,\kappa}} E_{k,\nu} \right] = \frac{1}{\sum_{\kappa=1}^{\nu} N_\kappa e_{0,\kappa}} \mathbb{E}[E_{k,\nu}] = \mathbf{E}_k. \tag{22}$$

Now, we consider the variance of the normalized iterative impulse response.

$$\mathbb{V}\left[\frac{1}{\sum_{\kappa=1}^{\nu} N_\kappa e_{0,\kappa}} E_{k,\nu} \right] = \frac{1}{\sum_{\kappa=1}^{\nu} N_\kappa^2 e_{0,\kappa}^2} \mathbb{V}[E_{k,\nu}] = \frac{1}{\sum_{\kappa=1}^{\nu} N_\kappa^2 e_{0,\kappa}^2} \sum_{\kappa=1}^{\nu} N_\kappa e_{0,\kappa}^2 V_k. \tag{23}$$

The underlying theory behind iterative ray tracing comes from Monte Carlo simulation theory, which states that the simulation precision increases as the number of samples, or rays, increases. However, this relies on the samples being identically distributed. Consequently, the initial energy should be constant for all rays, $e_{0,\kappa} = e_0$. Inserting this into Equation (23) gives

$$\mathbb{V}\left[\frac{1}{\sum_{\kappa=1}^{\nu} N_\kappa e_{0,\kappa}} E_{k,\nu} \right] = \frac{1}{\sum_{\kappa=1}^{\nu} N_\kappa} V_k. \tag{24}$$

Here, the variance decreases by the total number of rays in all iteration steps. This is equivalent to the behavior in the classical algorithm and is the suggested method for this algorithm.

In conclusion, if multiple ray tracing simulations are run independently for the same space using the same initial ray energy, the sum of their results normalized by the total emitted energy has the same expected value and variance as a single ray tracing simulation using the total number of rays emitted in all simulations.

4. Method

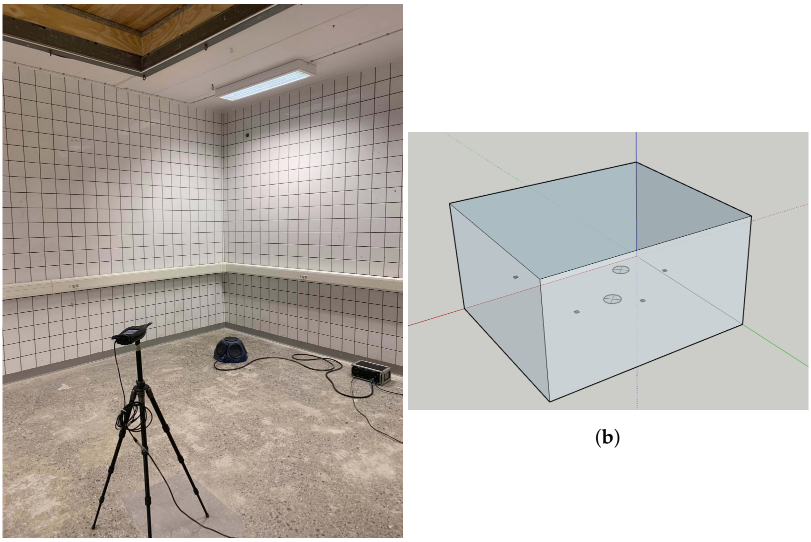
In this section, the methods used to test the iterative ray tracing algorithm and the estimated expected value and variance are described. As a first step, the space used for digital simulations and measurements is described. The measurements are described, as well as the simulation setups used. Finally, the implementation of the ray tracer, post-processing of the results, and some timing tests are commented.

4.1. Measurement and Simulation Space

The space used for measurements and simulations was the lower chamber in the impact sound lab at the Division of Engineering acoustics, Lund University (see Figure 2). The space has an approximate volume of 95 m^3 and a mid-frequency reverberation time of about 1.3 s. During the measurements, a suspended acoustic ceiling of porous absorbers was mounted in the center of the ceiling. This common acoustic treatment generally leads to a sound field with a net flow of acoustic energy towards the ceiling. This reduces the diffuseness in the sound field.

The space is rectangular, indicating that there should be strong and clear influences from eigenfrequencies in the low frequency range. Since ray tracers, in general, do not take resonance into account, the simulation results are expected to be relatively poor in the low frequency range. The Schroeder frequency can be used as a lower bound of the cut-off between these two cases, and it is found to be approximately 234 Hz [14]. In this project, especially the frequency bands centered on 125 Hz and 250 Hz are expected to be affected by resonances.

Acoustic impulse response measurements were obtained for validation of the ray tracer. Measurements were performed based on the international standards for room acoustic measurements [20] and have been previously described in 2021 [15]. An exponential sweep were used to obtain impulse response measurements. The measurements were performed using the open-source software REW version 5.19 (https://www.roomeqwizard.com/, accessed on 5 October 2022). A laptop with the software installed was connected to an Audio 8 DJ soundcard from Native instruments (Berlin, Germany) (https://www.native-instruments.com/en/, accessed on 5 October 2022), which connected to a Brüel & Kjær amplifier type 2734 (Hottinger Brüel & Kjær, Virum, Denmark) (https://www.bksv.com/en/transducers/acoustic/sound-sources/power-amplifier-2734, accessed on 29 August 2021). The amplifier fed the excitation signal to a dodecahedral loudspeaker. As a microphone, a Brüel & Kjær 2270 Sound Level Meter and Analyzer (Hottinger Brüel & Kjær, Virum, Denmark) (https://www.bksv.com/en/instruments/handheld/sound-level-meters/2270-series/type-2270-s, accessed on 5 October 2022) was used. A total of eight configurations were measured, two source positions and four listener positions, which are marked in Figure 2b. Due to technical and practical limitations, the sound source was located closer to the floor than what is prescribed in the standards.



(**a**) (**b**)

Figure 2. Photograph of a measurement setup (**a**) in the space measured in this project, although the suspended ceiling of acoustic absorbers was not installed at the time of this picture. Note that the speaker depicted is semi-dodecahedral, whereas the speaker actually used was dodecahedral. (**b**) shows the digital model of the same space with source (circular) and listener (small squares) positions marked.

4.2. Simulation Setups

Simulations were performed in the model space using both a classical ray tracer and using the iterative algorithm. Two main sets of simulations were performed, one to evaluate the theoretical estimates of the expected value and variance and one to evaluate the similarity between the classical and iterative algorithms. Both are described below.

When testing the estimates for expected value and variance, the surface absorption coefficient was constant across all frequency bands, but the air attenuation and scattering parameters were frequency dependent. The surface absorption was set uniformly to $\alpha = 0.1$ across all surfaces and frequencies and air attenuation and scattering parameters are presented in Table 1. Air attenuation is set based on the environmental conditions of the measurement space during measurements, and the scattering parameters were set based on the results of earlier research [15]. Since the scattering parameter affects how diffused the simulated sound field is, their variations across the frequency range should lead to variations where frequency bands with higher scattering coefficients should match the theoretical results better.

Table 1. Simulation parameters described for each octave band by frequency (f). s is the dimensionless surface scattering parameter, constant for all surfaces in the space. γ is the air attenuation parameter, defined in m^{-1}. Finally, the dimensionless absorption coefficients α are given for the surfaces in the model. Note that the surface absorption parameters were not used in the simulations designed to evaluate the estimates of expected value and variance. Surface scattering and absorption parameters are based on calibration from earlier research [15].

f(Hz)	s	γ (m^{-1})	α, Walls	α, Floor	α, Ceiling Absorbers
125	0.01	0	0.0734	0.0596	0.25
250	0.13	0	0.056	0.0453	0.80
500	0.05	0.00064	0.0453	0.0453	0.95
1000	0.13	0.0015	0.04	0.045	0.95
2000	0.1	0.00208	0.045	0.045	0.95
4000	0.25	0.00532	0.045	0.045	0.95

Simulations were also used to examine the difference between the iterative and the classical ray tracing algorithms. In these simulations, surface absorption parameters varied across surfaces and frequency bands, shown in Table 1. Air attenuation and surface scattering parameters were identical to what was used in simulations for theory evaluation.

For both of the configurations described above, ray tracing simulations were performed using both the classical and the iterative algorithms. A total of eight configurations were simulated, two source positions and four listener positions (as shown in Figure 2b). In the classical ray tracer, 900,000 rays were used. In the iterative algorithm, each iteration used 9000 rays and there were 100 iterations. This results in a total of 900,000 rays contributing to the final impulse response, the same number as in the classical algorithm. A total of 15 simulations were performed for each configuration, each yielding an energetic impulse response across six octave bands in the range 125 Hz–4 kHz.

4.3. Ray Tracing Implementation

An implementation of a classical ray tracer was used for the simulations. The NVIDIA OptiX ray tracing engine version 4.1.1 [21] is used as a foundation for the program. The OptiX engine traces the paths of traveling rays and detects collisions with the digital model. Programs to generate rays, to determine the direction and energy of rays after reflection, and to record ray collisions with the detector are implemented in CUDA to run on the graphics card. These are accessed through a C++ wrapper, which operates on the host computer. The simulations are performed on an NVIDIA GeForce GTX 1070 on a machine with an Intel i5 3.8 GHz processor with four cores.

In the program, new rays are randomly generated from a uniform spherical distribution at the source. Scattering is implemented by randomly determining the new direction of rays as either the direction predicted by Schnell's law (i.e., the specular direction) or a direction randomly generated from an ideal diffused distribution. This implementation is arguably the most commonly used implementation [1,5] and found to outperform the most common alternatives [15].

As shown in Figure 1 and Section 3.2, the iterative algorithm can be implemented in addition to, and separately from, a standard ray tracer, as it acts only on the produced energetic impulse responses. In this study, the iterative algorithm is only partially implemented, as only the step for addition of impulse responses is tested. Pseudo code for the implementation in question is shown in Listing 1. For the iterative algorithm, all rays have an initial energy of 1. The same number of rays are emitted in each iteration step, and the total number of rays emitted is recorded and updated after each simulation. Since the initial ray energy is 1, the total number of rays used coincides with the normalization factor that should be used to ensure that the iterative algorithm has the correct expected value. As new impulse responses are calculated, they are added to the current impulse response in the C++ wrapper. The impulse response is only normalized when it is passed on to be used externally.

Listing 1. Pseudo code for the partial iterative implementation.

```
int number_of_rays;
float normalization_factor;
Raytracer raytracer;
float[] new_impulse_response, iterative_impulse_response;

while (true) {
    new_impulse_response = raytracer.run(number_of_rays);
    iterative_impulse_response += new_impulse_response;
    normalization_factor += number_of_rays;
    }
return iterative_impulse_response/normalization_factor;
```

The raw impulse responses have a sampling rate of 44.1 kHz and are 8 s long.

4.4. Post-Processing of the Results

In order to interpret the results of the simulations, some post-processing was necessary. In particular, graphical comparisons of impulse response data were found to require a low temporal resolution to reduce the noise. Consequently, the impulse responses used in plots have a sampling rate of 441 Hz, giving a time resolution of about 2.23 ms. These data were obtained by summing the energy in the 100 time steps of the original impulse response that would be overlapped by the extended time period. Since ray tracers produce energetic impulse responses, this can be achieved.

Room acoustic variables were extracted from the simulated and measured impulse responses. Reverberation time (T_{20}), early decay time (EDT), and speech clarity (C_{50}) were extracted from the measured impulse responses using ITA-toolbox for MATLAB version R2020b, developed in Aachen [22], and the same room acoustic parameters were extracted from the simulated impulse responses using a MATLAB script.

4.5. Timing Tests

Finally, the time requirements for the classical algorithm and the iterative algorithm were compared. For the ray tracer used in this study, each ray tracing step can be divided into two parts. The first is the ray tracing procedure itself, complete with ray generation, path calculation, and energy evaluation, where an energetic impulse response is formed on the graphics card. The second part of the ray tracing step is the retrieval of data from the graphics card to the working memory of the host computer. Only the first step depends on the number of rays outright, and the second is independent of the number of rays. In iterative ray tracing, the retrieval of data occurs more frequently, and the overall time consumption per ray, thus, increases. This effect is exacerbated if the number of rays in each update step is very small.

The number of rays used generally has a dominating impact on the overall simulation time and should, therefore, be chosen carefully. For the classical algorithm, it has been determined that 540,000 rays is sufficient to accurately model the reverberation time, early decay time, and speech clarity [15]. It is used for simulation time measurements in this study.

For the iterative algorithm, the number of rays depend on the time available for simulation, on the order of 50 ms for real-time performance [12]. However, that number of rays should still provide a reasonable model of the acoustic field in the space. A possible limit may be, for example, that the simulated impulse response should have room acoustic parameters within 1 JND of the actual value (as predicted by the GA model), 75% of the simulations. For the ray tracer used in this project, further optimization is needed before the simulation results are reasonably accurate after 50 ms simulation time. Instead, the proposed limit of 75% of the studied room acoustic parameters, being within 1 JND of the actual value, is used. This corresponds to 81,000 rays.

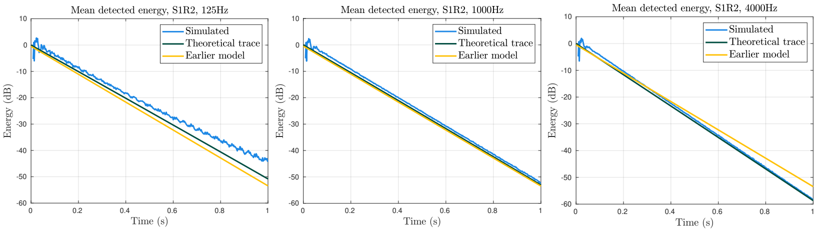
5. Results

5.1. Validation of the Theoretical Model

The theoretical expressions for the expected value and the variance of a ray tracing simulation (derived in Section 3.1) were compared to estimates formed from test simulations and to an earlier theoretical model described in [14]. In particular, whether the effects of varying the number of rays is similar for all data sets was examined. If it is, it shows that the theoretically derived estimates describe important behavior in the real simulations.

The simulated estimates of the expected value and variance were derived using standard methods, from the 15 sets of simulations described in Section 4.2. Each source and listener configuration was treated separately, but no systematic or significant differences were found when comparing to the theoretically-derived estimates. The full set of graphs is available as supplementary material.

In Figure 3, the expected value estimated both from simulations and theoretical expressions is compared. As expected, the theoretical expressions are less accurate when the sound field is less diffused, in the early parts of the impulse response and in cases with low scattering. In these cases, the theoretical estimates tend to underestimate the total detected power. Since diffused sound fields generally tend to decay faster and have lower energy overall, this supports the theory that the simulated sound field is not diffused in these cases.



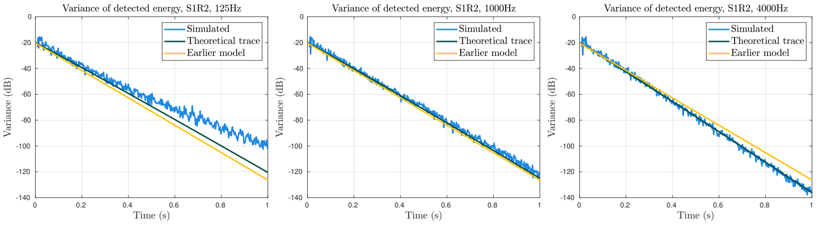
(**a**) (**b**) (**c**)

Figure 3. Comparison between the theoretical estimate of the mean detected energy (Equation (7), green), the estimate found in [14] (yellow), and an approximation based on simulations (light blue). Note that only the scattering coefficient and air absorption vary between the octave bands. For low scattering, (**a**), the simulation results exhibit a ringing behavior which is not replicated in either theoretical estimate. In (**b**), both theoretical estimates match the simulation well. In (**c**), the air attenuation affects the results, and the new theoretical estimate outperforms the simplified expression.

The two theoretical expressions are very similar, especially when air attenuation has no significant impact. When air attenuation affects the results, the newly derived estimate is significantly better. This is not surprising, and it is worth mentioning that air absorption can very easily be added also to the older estimate. When air absorption is not significant, the new estimate of the expected value is slightly smaller, as compared to the previous estimate. Since the earlier estimate makes an even stricter assumption on the diffuseness of the sound field (as discussed in Section 3.1), this can be interpreted as the earlier expression describing a more diffused sound field, which consequently has lower energy.

In Figure 3a, the expected detected energy for the octave band centered on 125 Hz is shown. In the results from simulations, a ringing behavior is seen. This is likely due to rays following a cyclical path through the simulated space, hitting the detector once every lap. Since all rays are emitted at the same time, $t = 0$, all rays following the same path hit the detector at the same time. This behavior can be seen in ray tracers if the simulated space have very low scattering. It is not related to acoustic resonance.

In Figure 4, the theoretical estimates of the variance are shown, together with an estimate from 15 simulations. Although the variance estimate derived from simulations is more noisy, its general trend (when the sound field is sufficiently diffuse) is matched very well by the theoretical expressions. The influence of simulation noise on the estimate of the variance is expected to be larger than its influence on the expected value, since the variance is approximately squared, compared to the measurement data itself. This amplifies all noise.

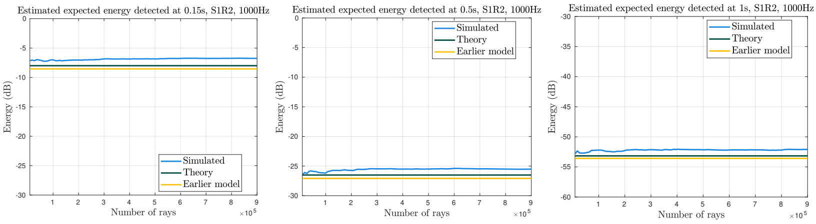


(**a**) (**b**) (**c**)

Figure 4. Comparison of the estimated variance of the detected energy in the impulse response, using the new expression Equation (7, green), the estimate found in [14] (yellow) and the results, based on 15 simulations (light blue). (**a**) shows the result for low scattering, (**b**) the results with more scattering and (**c**) the results when both scattering and air attenuation affects the simulation.

The results found for the expected value are mostly replicated in the results for the variance. That is, the theoretical expressions are more accurate for higher scattering and not accurate for the very early impulse response. In addition, it is found that the simplified expression and the expression derived in this paper are similar. When air attenuation affects the simulation, the expression which takes this into account is significantly better.

In Figure 5, the expected energy for each time is shown for increasing number of rays. Both the theoretical expressions are independent of the number of rays and are shown as horizontal lines in the plots. The estimated results do not show any dependence on the number of rays. As noted in Section 1, the expected value of Monte Carlo simulations should be independent of the number of samples, and any result other than what was found here would be alarming.

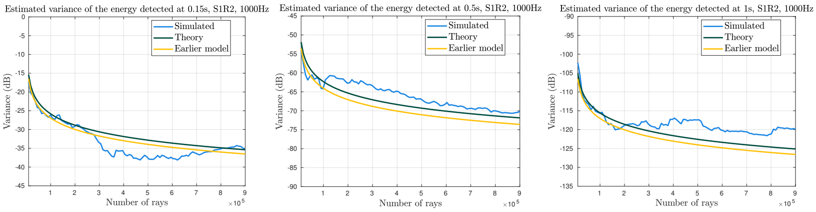


(**a**) (**b**) (**c**)

Figure 5. Graphs illustrating how the expected value of the impulse response at 0.15 s (**a**), 0.5 s (**b**) and 1 s (**c**), estimated theoretically and from simulation data, changes as the number of rays change. The theoretical estimates are both constant, and this behavior is generally repeated by the simulation data.

In the three graphs in Figure 5, the simulation results are slightly larger than the theoretical estimates (as was found in Figure 3), possibly due to the level of diffuseness in the simulations not reaching the level of diffuseness assumed in the theoretical expressions. However, the difference between the simulated result and the theoretically calculated results does not seem to change as time grows from 0.15 s to 1 s. A possible explanation is that the sound field after 0.15 s has reached its maximum level of diffuseness.

The variance, as it depends on the number of rays, is shown in Figure 6. As noted in Section 3.1, the variance decreases as the number of rays increase. The rate of change seems similar for all three graphs shown, but the computation noise in the simulated data is too large to make more precise statements.



(**a**) (**b**) (**c**)

Figure 6. Graphs illustrating how the estimated and modelled variance in a given time frame changes as the number of rays used changes, at 0.15 s (**a**), 0.5 s (**b**) and 1 s (**c**). The two th eoretical estimates decrease as the number of rays increase and seem to do so at a similar or equal rate. While the results from simulations are more noisy, they seem to follow a similar pattern and are quite close to both theoretical estimates.

5.2. *Validation of the Algorithm*

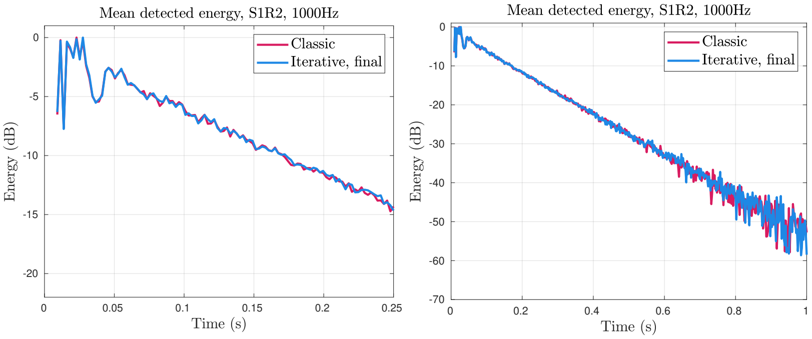
In this section, the results from the classical and the iterative algorithms are compared. In particular, it is examined whether the precision and accuracy of the two algorithms are equivalent for a large number of rays. This is primarily evaluated by comparing the estimated expected value and variance of the impulse responses produced using the maximum number of rays. Some room acoustic parameters are also studied to help determine whether the two algorithms produce perceptually similar responses. The room acoustic parameters are shown, together with measurement data for the space in question, to put the results into context. For a more detailed discussion of the correspondence between measurements and simulations, the reader is referred to [15].

As an initial test, the total detected energy in the broadband impulse response (simply the sum of all the energy detected in each simulation) was compared using a Mann–Whitney U test (Wilcoxon rank sum test) [23,24]. This is a non-parametric test on two data sets that indicates the likelihood that these or more dissimilar measurements could come from identical populations. As such, low values on the test signifies a higher likelihood that there is a difference between the two data sets. In this case, no significant difference could be found between the classical and the iterative algorithms. The results are shown in Table 2 for each source and listener position separately, since the impulse responses for each source and listener position are unique.

Table 2. p-Value from a Mann–Whitney U test comparing the total energy detected in the impulse response for the iterative and classical algorithms. The results do not show a significant difference between the two data sets for any of the positions.

	S1R1	**S1R2**	**S1R3**	**S1R4**	**S2R1**	**S2R2**	**S2R3**	**S2R4**
p-Value	0.23	1.00	0.74	0.15	0.13	0.71	0.19	0.37

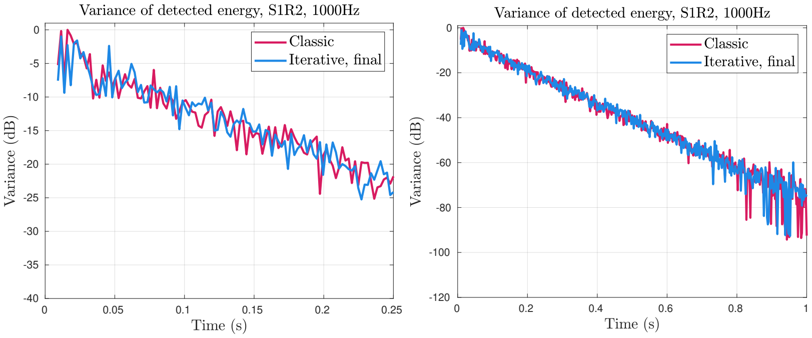
The mean detected energy in the broadband impulse response was compared for the two algorithms, and some sample graphs are shown in Figure 7. The full set of graphs are available in the supplementary material. No significant differences can be seen. In particular, the results for the very early parts, containing the vital early reflections, are almost identical. Towards the end of the response, some discrepancies can be seen, but these are small in absolute terms and likely due to simulation noise. The similarities in the early response and in the decay rate show that the estimated expected values are very similar and likely at least perceptually identical.



(**a**) (**b**)

Figure 7. The mean detected energy in the impulse response for the classical and iterative algorithms, the first 0.25 s of the impulse response in (**a**) and the first 1 s (**b**). The results are very similar.

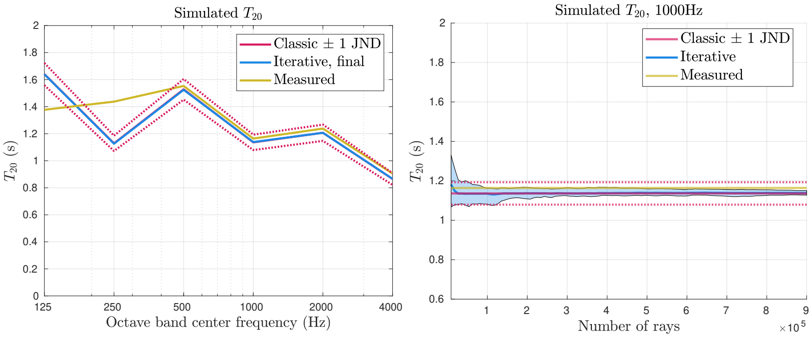
In addition to comparing the mean of the simulations, the estimated variance was compared, as shown in Figure 8 and in the supplementary material. The estimates of the variances are more noisy, similarly to the variance found when checking accuracy of the theoretical estimates. It seems that the variances of the two algorithms have an equal decay rate over the impulse response.



(**a**) (**b**)

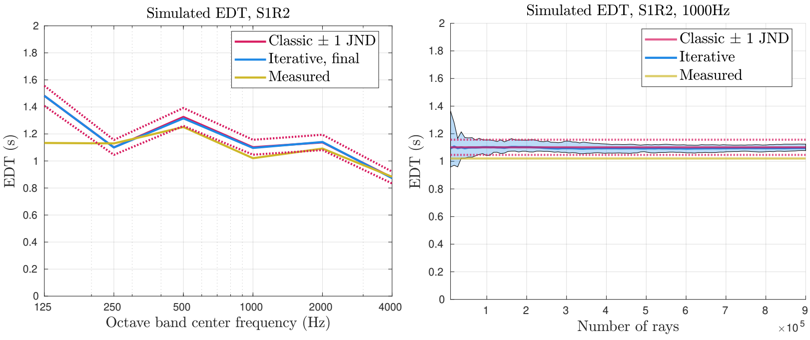
Figure 8. The variance of the detected energy in the impulse response for the classical and iterative algorithms, the first 0.25 s of the impulse response (**a**) and the first 1 s (**b**). The results are quite noisy, but similar.

In addition to comparison of the broadband energetic impulse response, the reverberation time T_{20}, early decay time (EDT), and speech clarity C_{50} were studied. Some results are shown in Figures 9a–11a. In these graphs, the averages from 15 simulation runs are shown for the maximum number of rays. The reverberation time T_{20} has been spatially averaged, in accordance with [20]. No significant difference can be seen between the estimated expected value of the two algorithms for any of the room acoustic parameters.



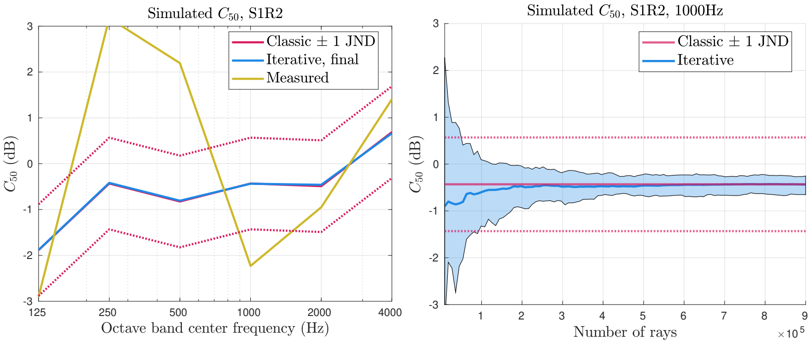
(**a**) (**b**)

Figure 9. Reverberation time T_{20} measured and obtained from simulations using the classic and the iterative algorithm. Dashed lines indicate values 1 JND away from the results from the classic ray tracer. In (**a**), the results for the maximum number of rays are shown. (**b**) shows the estimate for 1000 Hz using an increasing number of rays in the iterative algorithm, as well as the results for the full classic simulation. The shaded area illustrates the spread of estimates in the iterative algorithm.



(**a**) (**b**)

Figure 10. Early decay time EDT for one position, measured and estimated from classic and iterative simulation data sets. In (**a**), the results for the maximum number of rays are shown. (**b**) shows the estimate for 1000 Hz increasing the number of rays in the iterative algorithm and showing the results for the full classic simulation. The shaded area illustrates the spread of estimates in the iterative algorithm.



(**a**) (**b**)

Figure 11. Speech clarity C_{50} for one position, measured and estimated from classic and iterative simulation data sets. In (**a**), the results for the maximum number of rays are shown. (**b**) shows the estimate for 1000 Hz increasing number of rays in the iterative algorithm and shows the results for the full classic simulation. The shaded area illustrates the spread of estimates in the iterative algorithm.

For the iterative algorithm, the room acoustic parameters were also studied, as the number of steps and, consequently, the number of rays increased, as shown in Figures 9b–11b. Based on the calculations in Section 3.1, the expected value of the impulse response should be constant while the variance decreases. This should translate to a decrease in variance of the room acoustic parameters as the number of steps increase, but it does not necessarily mean that their expected value is constant.

In the graphs shown in Figures 9b–11b, the room acoustic parameters for the octave band centered on 1000 Hz are shown as the number of steps in the iterative algorithm increases. The average from 15 simulations is shown as thick lines, and the area between the minimum and maximum parameter values is shaded. This illustrates the spread of the estimates over the whole set of simulation runs. As expected, the variations decreases as the number of steps increase, and when the total number of rays reaches about 100,000 (about 11 steps), the variations are about 1 JND.

It can also be noted that the estimated expected values for reverberation time T_{20} and C_{50} are not constant for all steps in the iterative algorithm. It is possible that this reflects a real effect for small numbers of rays, where the increased variability in the impulse response leads to a systematic difference in the calculated room acoustic parameters. It is also possible that the increased variability leads to a situation where the sample size of 15 is not large enough to ensure that the expected value is estimated accurately. However, the estimated expected values does not look particularly noisy, even for small numbers of steps, so it seems more likely that there is a systematic difference, due to how the room acoustic parameters are calculated.

As discussed in Section 4, the room acoustic parameters were used as a tool to estimate an appropriate number of rays to use in the iterative algorithm. The chosen number of rays is 81,000, and it can be seen in Figures 9–11 that, for this number of rays, the variations of the room acoustic parameters are about 1 JND for this frequency band.

5.3. Computation Time

Iterative ray tracing is expected to improve the performance of classical ray tracing by adapting to the speed or precision as necessary. The results presented so far focus on showing that the iterative algorithm is as precise as the classical ray tracing algorithm, given sufficient time. In this section, the performance, in terms of speed, is evaluated. One aspect that is considered is how the time consumption per ray changes when iterative ray tracing is implemented. Another aspect is how large the overall time improvement is when the number of rays is decreased, as this indicates if it is possible, in general, to come closer to real-time performance by using fewer rays. As mentioned in Section 4, the ray tracer used in this study is not fully optimized for speed and cannot achieve real-time performance. The number of rays to use in each step of the iterative algorithm was consequently determined using other parameters and set to 81,000. For the classical ray tracing step, the full number of rays used was 540,000. Based on earlier research, this is sufficient for the space in question [15].

The simulation time for each algorithm is shown in Table 3. The ray tracing average is the average time for a full ray tracing step, yielding energetic impulse responses for six octave bands and four listener positions. The retrieval average details the time consumption for transferring the impulse response data from the graphics card to the host computer (consisting of twelve impulse responses 8 seconds in length with a sample rate of 44.1 kHz). The update trace takes about 15% of the time required by the full simulation. The number of rays used in the iterative algorithm is also 15% of the number of rays used in the full simulation, which suggests that the time consumption is proportional to the number of rays, at least in this implementation of ray tracing. It is also noted that the data retrieval step takes longer for the iterative algorithm. This is due to the additional step of adding the calculated impulse response to that which has been calculated in previous time steps. The increase in total time is short, however.

Table 3. Comparison of the time consumption for the classical algorithm and for the iterative algorithm. Computation time is presented for the ray tracing simulations and the transfer of simulation results from the graphics card to the host computer.

Algorithm	Ray Tracing Average ± std (s)	Transfer Average ± std (s)
Classic, 540,000 rays	2.055 ± 0.039	0.0250 ± 0.0012
Iterative, 81,000 rays	0.307 ± 0.003	0.0260 ± 0.0004

The iterative algorithm reaches and exceeds the number of rays in the classical algorithm after seven steps with the current configuration. Assuming that the total time of a simulation is adequately estimated as the sum of the two steps detailed above, seven runs would take approximately 2.33 s, or about 0.3 s more than a full classical trace. The total number of rays in the classical trace is then exceeded by 27,000 rays. In total, the time consumption per ray is approximately 7% higher in the iterative algorithm in this specific configuration, and the approximate time consumption per ray is 4 µs for both algorithms.

6. Discussion

In this paper, an algorithm that improves the adaptability of ray tracing simulations in real-time applications has been presented. When the simulation time limits imposed by real-time performance negatively affects the simulation precision, iterative ray tracing can be used to improve simulation precision over time, eventually reaching optimal ray tracing performance. It is shown that the simulations produced by the iterative algorithm are equivalent to those produced using the classical algorithm with the same total number of rays, by both revising the theoretical estimates of the expected value and variance and comparing results from the two algorithms outright.

Additionally, a theoretical expression is derived for the expected outcome and variance of a ray tracing simulation in a diffused sound field. These expressions can be used to increase understanding of ray tracing simulations and their variability, which can help in determining what number of rays to use and what factors affect the reliability of ray tracing simulations. In addition, since ray tracing simulations are an unbiased tool for calculating the energetic impulse response predicted by a geometrical acoustics model, the estimated expected value can be used explicitly to find the impulse response of a space. This is only accurate if the sound field is diffused, and in real spaces, the sound field is practically never diffused. However, the later parts of the sound field, in some cases, can be approximately diffused. The expression could, thus, be an acceptable way to estimate the so-called reverberating tail of the impulse response.

The calculated expected value and variance are very similar to previously derived estimates, while being rather more complex to evaluate. The increased complexity is introduced as both the ray paths and the ray energy are considered stochastic, and the expressions in this paper can, consequently, be said to be more stringent, at least in that aspect. In that sense, the presented calculations show that the simplifications made in earlier derivations are appropriate at least in some spaces. The results presented in this paper also emphasizes that air absorption should be considered.

The iterative ray tracing algorithm has a similar performance as the classical algorithm if the total number of rays is equal. This is shown using the theoretical expressions for expected value and variance by examining the room acoustic parameters and by outright comparing the average impulse responses from the two algorithms.

In general, the time consumption per ray is greater for the iterative algorithm. However, the difference is small. It is not likely that this effect, even in real-time applications, would be sufficiently large to have a significant impact on how many rays can be used in each simulation step.

Iterative ray tracing needs to be tested further before it can be deployed in real applications. Importantly, it is not yet clear whether the algorithm can be used to speed up the simulation without perceptual degradation. This needs to be tested by listening tests with users, preferably in full virtual reality. Furthermore, the resolution for the spatial discretization needs to be determined, although this can likely be estimated quite well from the existing research on acoustic virtual reality.

The future development of the algorithm could be focused on the spatial aspects. For example, linear interpolation is mathematically equivalent to the operations performed when combining the impulse responses from different runs. If the impulse responses are from different locations, the summation and normalization step gives a spatial interpolation of the various positions, and modifications of the normalization value can be used to change the relative position of the interpolated impulse response. This could be used to modify the spatial resolution and to counteract any issues introduced by it. It may also be possible to remove the need for an explicit spatial resolution by, instead, using a moving average process, where the current impulse response consists of a linear combination of several of the latest calculations. This would implicitly perform a spatial interpolation and potentially also mitigate the risks of perceptual degradation when a new impulse response is to be calculated.

An advantage to iterative ray tracing is that it can be used well in combination with other methods of improving ray tracer performance, in terms of either accuracy or speed. The algorithm requires that the results from different runs of the ray tracer are statistically independent, but it does not have any specific limitations on how the ray paths are calculated and is very flexible, in terms of what algorithms and models it can be combined with.

7. Conclusions

Iterative ray tracing improves the overall precision of ray tracing simulations when real-time restrictions may otherwise adversely affect simulation performance. In comparison to classic ray tracing, it is as fast, but can also iteratively improve the simulated impulse response over time. The algorithm is flexible enough to be used in conjunction with other methods of improving ray tracer performance, both in terms of speed and in terms of accuracy.

In this study, and for this test space, it was shown that using as few as 15% of the rays produced simulation results was accurate to within 1 JND of the studied room acoustic parameters (T_{20}, EDT, and C_{50}) approximately 75% of the time. This is seven times faster than what is needed for full accuracy. It has also been shown that iterative ray tracing can take advantage of this significant improvement in speed, while still providing full accuracy as quickly as the classical algorithm.

The next step in the development of the iterative ray tracing algorithm should be to determine an adequate spatial resolution and then implementing the iterative algorithm, in conjunction with a fast ray tracing step and a real-time auralization engine.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/acoustics5010019/s1, Supplementary graphs showing results for additional octave frequency bands and source and listener positions.

Author Contributions: Conceptualization, H.A.; methodology, H.A. and N.-G.V.; software, H.A.; validation, H.A.; formal analysis, H.A.; writing—original draft preparation, H.A.; writing—review and editing, H.A., N.-G.V. and D.B.H.; visualization, H.A.; supervision, N.-G.V. and D.B.H.; funding acquisition, D.B.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Swedish Research Council, grant number 2016-01784. The APC was partially funded by Lund University.

Data Availability Statement: Raw data and source code are available upon request.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Cox, T.J.; D'Antonio, P. *Acoustic Absorbers and Diffusers*, 3rd ed.; Taylor & Francis Group: Boca Raton, FL, USA, 2017.
2. ODEON Room Acoustics Software. *Simulation Software*; ODEON A/S:Lyngby, Denmark, 2020.
3. Dalenbäck, B.I. *CATT-Acoustic v9.1 Powered by TUCT v2*; Simulation Software: Gothenburg, Sweden, 2021.
4. van der Harten, A. Pachyderm Acoustical Simulation: Towards Open-Source Sound Analysis. *Archit. Des.* **2013**, *83*, 138–139. [CrossRef]
5. Savioja, L.; Svensson, P. Overview of geometrical room acoustic modeling techniques. *J. Acoust. Soc. Am.* **2015**, *138*, 708–730. [CrossRef] [PubMed]
6. Kuttruff, H. *Room Acoustics*, 6th ed.; CRC Press: Boca Raton, FL, USA, 2017.
7. Savioja, L.; Manocha, D.; Lin, M.C. Use of GPUs in room acoustic modeling and auralization. In Proceedings of the International Symposium on Room Acoustics, ISRA 2010, Melbourne, VIC, Australia, 29–31 August 2010.
8. NVIDIA. *VRWorks Audio*; Computer Software; NVIDIA: Santa Clara, CA, USA, 2018. Available online: https://developer.nvidia.com/vrworks/vrworks-audio (accessed on 13 March 2023).
9. Mehra, R.; Rungta, A.; Golas, A.; Lin, M.; Manocha, D. WAVE: Interactive Wave-based Sound Propagation for Virtual Environments. *IEEE Trans. Vis. Comput. Graph.* **2015**, *21*, 434–442. [CrossRef]
10. Treble Technologies. *Treble Technologies*; Computer Software: Reykjavik, Iceland, 2022.

11. Selfridge, R.; Cook, J.; McAlpine, K.; Newton, M. Creating Historic Spaces in Virtual Reality using off-the-shelf Audio Plugins. In Proceedings of the Conference on Immersive and Interactive Audio 2019, New York, NY, USA, 27–29 March 2019.
12. Schröder, D. Physically Based Real-Time Auralization of Interactive Virtual Environments. Ph.D. Thesis, Institute of Technical Acousitcs, Aachen, Germany, 2011.
13. Poirier-Quinot, D.; Katz, B.; Noisternig, M. EVERTims: Open source framework for real-time auralization in VR. In Proceedings of the AM'17, London, UK, 23–26 August 2017; ACM: New York, NY, USA; London, UK, 2017.
14. Vorländer, M. *Auralization. Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 1–335. [CrossRef]
15. Autio, H.; Vardaxis, N.G.; Bard Hagberg, D. The Influence of Different Scattering Algorithms on Room Acoustic Simulations in Rectangular Rooms. *Buildings* **2021**, *11*, 414. [CrossRef]
16. Graham, C.; Talay, D. *Stochastic Simulation and Monte Carlo Methods*; Springer: Berlin/Heidelberg, Germany, 2013.
17. Autio, H.; Bard Hagberg, D. An Attention-Guided Algorithm for Improving the Performance of Acoustic Simulations. In Proceedings of the 23rd International Congress on Acoustics: Integrating 4th EAA Euroregio 2019, Aachen, Germany, 9–13 September 2019; Ochmann, M., Vorländer, M., Fels, J., Eds.; Deutsche Gesellschaft für Akustik: Aachen, Germany, 2019; pp. 2619–2626.
18. Schneiderwind, C.; Neidhardt, A. Perceptual differences of position dependent room acoustics in a small conference room. In Proceedings of the International Symposium on Room Acoustics, Amsterdam, The Netherlands, 15–17 September 2019; Lautenbach, M., Witew, I.B., de Vries, D., Hornikx, M., Eds.; ISRA: Amsterdam, The Netherlands, 2019; pp. 499–506.
19. Shinn-Cunningham, B.; Ram, S. Identifying where you are in a room: Sensitivity to room acoustics. In Proceedings of the 2003 International Conference on Auditory Display, Boston, MA, USA, 6–9 July 2003; pp. 21–24.
20. *ISO 3382-1:2009*; Acoustics-Measurement of Room Acoustic Parameters-Part 1: Performance Spaces. Technical Report; ISO: Geneva, Switzerland, 2009.
21. Parker, S.G.; Bigler, J.; Dietrich, A.; Friedrich, H.; Hoberock, J.; Luebke, D.; McAllister, D.; McGuire, M.; Morley, K.; Robison, A.; et al. OptiX: A general purpose ray tracing engine. *ACM Trans. Graph.* **2010**, *29*, 1–13. [CrossRef]
22. Berzborn, M.; Bomhardt, R.; Klein, J.; Richter, J-G.; Vorländer, M. The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic Measurements and Signal Processing. In Proceedings of the 43rd Annual German Congress on Acoustics, Kiel, Germany, 6–9 March 2017 .
23. Mann, H.B.; Whitney, D.R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *Ann. Math. Stat.* **1947**, *18*, 50–60. [CrossRef]
24. Wilcoxon, F. Individual Comparisons by Ranking Methods. *Biom. Bull.* **1945**, *1*, 80. [CrossRef]

Paper E

A statistical method for parameter estimation from Schroeder decay curves

Hanna Autio[a]
Delphine Bard[b]
Lund University, Division of Acoustics
Lund, Sweden

Abstract

The sound decay curve, measured by either the interrupted noise method or the Schroeder method, can provide relevant information pertaining to a room's acoustical properties. However, the relevant information is not always readily extractable from the measurements. In this study, measurements of the decay curves of the historically important Vadstena Abbey church in Sweden have been obtained using both of the classical methods. The test building was in normal use with high levels of background noise during the measurements. A line-fit analysis showed a trend where the Schroeder decay curves yielded a longer estimated reverberation time compared to the interrupted noise method for low frequencies. This indicates that the measurement results are affected by background noise. A statistical model was developed for the Schroeder decay curve in such circumstances. Using it, the maximum likelihood estimates for the deterministic decay parameters can be found even when there are multi-slope decay patterns or significant background noise. The model is straight-forward to implement and can find an estimate for a single-slope decay also when some commercial software fails. The performance for double-slope decay models, as well as a comparison to the results from the interrupted noise method, are analyzed.

1 INTRODUCTION

The reverberation time is one of the most used values to describe the acoustical properties of a space, and it is often used when establishing policies or evaluating existing spaces. The simplicity in measuring it makes for an appealing measure, and the properties it describes are intuitive and have a clear impact on the acoustic comfort.

---

[a] hanna.autio@construction.lth.se
[b] delphine.bard@construction.lth.se

Presented at Internoise 2018 Congress, August 2018

The reverberation time relates closely to the sound pressure decay in a space. When a source continuously excites the sound field in a space and then suddenly is turned off, the sound pressure will not decrease to zero immediately, but will decay over time. The reverberation time is defined as the time it takes for the sound pressure level to decrease by 60dB under these circumstances[2], and is a property of the room and varies by frequency.

There are two well-established ways of measuring the reverberation time. One is the interrupted noise method, which falls back onto the definition of the reverberation time itself and relies on fully exciting the sound field in the space and explicitly measuring the decay time. The other is the integrated impulse response method, which relies on Schroeder's integration method from [cite year]. This method measures the impulse response of the room, and uses a backwards integration method to estimate the sound decay. If the recorded impulse response is accurate, this provides an exact result without the need to measure multiple times. However, due to the backwards integration, the method is somewhat sensitive to disturbances caused by background noise.

The sound pressure decay contains a significant amount of information on the audial experience in a space, and the reverberation time is an exact descriptor of the sound energy decay only in some cases. Its derivation relies on a uniform distribution of sound energy, that is, a diffuse sound field. Where the sound field is not diffuse, variations in space (such that the sound decay is not constant over the full volume) can occur, and the sound decay curve may be curved such that a single value does not describe the behavior fully for both early and late decay. Non-diffuse sound fields can be found in coupled spaces (such as concert halls with a delimited stage), rooms where some, but not all, surfaces are absorbent, or large rooms in general[2].

As it is commonly assumed that the sound pressure level decays linearly, a linear regression estimation method is often used to determine the reverberation time. An alternative has been presented by Xiang[3] for Schroeder decay curves, and the methods developed in that project provides a method for evaluating the reverberation time as well as sound decay parameters for spaces with a curved decay.

Based on the considerations presented above, especially concerning a variety of models for the sound pressure decay, a probabilistic model applicable for many different decay patterns was developed. The log-likelihood function relies on an estimate of the integrated background noise rather than the model chosen for the sound pressure decay, and can be quite readily adapted to other decay patterns.

The model presented in this paper is found to provide estimates for the reverberation time and other sound decay parameters for an exponential sound decay model. It is also found that the optimization algorithm used as well as initial values might have a significant influence on the result.

2 MODEL

This section describes and establishes a probabilistic model for the Schroeder decay curve, based on an assumption that the measurements consist of two distinct parts; one deterministic and one random. The influence of the random background noise on the measurements is examined, and a model is developed that takes into consideration the mathematical operations used in the formulation of the Schroeder decay curve. A log-likelihood function that depends only on estimates of the random parts of the Schroeder decay curve is defined, and so the function can be used to find maximum likelihood-estimates of parameters for different models of the deterministic decay.

2.1 Assumptions on the deterministic sound pressure decay

While the method and the log-likelihood function is valid for various models of the deterministic sound decay, only the classical exponential decay patterns have been examined and are discussed in this paper.

Furthermore, it is assumed that in the case of a curved decay, the decaying sound field can be described by two (or more) distinct exponentially decaying terms. This implies that, at least mathematically, the sound field can be divided into discrete segments where there is no energy transfer between the segments. This model corresponds to previous suggestions[4;1]. In doing so, we assume that the ratio of the total energy within each segment is determined by the space itself. That is, if the energy within (for example) two different sound fields can be described by A_1 and A_2 at time 0, the ratio $\frac{A_1}{A_2}$ at that time is constant for each measurement configuration, although the absolute energy may not be.

2.2 A statistical model for the Schroeder decay curve

After an impulsive excitation of the sound field, the sound pressure field is assumed to consist of two parts, the impulse response $h(t)$ and some background noise $e(t)$, which are here regarded as discrete functions of the time since the impulsive input. These yield a measurement of the sound pressure at time $t = t_k$ as

$$\xi\left(t_k\right) = h(t_k) + e(t_k). \tag{1}$$

The Schroeder decay curve $d(t_k)$ is formed from these measurements as follows:

$$d\left(t_k\right) = \sum_{j=k}^{K} \left(\xi\left(t_j\right)\right)^2 = \sum_{j=k}^{K} \left(h^2(t_j) + 2h(t_j)e(t_j) + e^2(t_j)\right), \tag{2}$$

where K is the index of the last measured time point.

In the ideal case where there is no background noise, the only non-zero term in this equation corresponds to the squared impulse response, and fully describes the sound level decay.

The cross-term containing both information from the impulse response and the background noise will (approximately) vanish in circumstances where the two signals are uncorrelated and at least one of them is zero-mean. As both signals are pressure variations around a mean, they can be assumed to be zero-mean. However, as the background noise can be described as the convolution between some input noise and the impulse response of the space in which we measure, they are likely not independent. Nevertheless, the cross-term will be neglected and it is assumed that any errors introduced by this will be small compared to the variations introduced by the squared background noise.

According to classical methods, the sound pressure decay can be described by an exponential curve. For spaces with curved decays, there can be several exponential terms contributing to the decays pattern. This model lets us describe the Schroeder decay curve as

$$d(t_k) = \sum_{s=1}^{S} A_s \exp\left[-T_s t_k\right] + \sum_{j=k}^{K} e^2(t_j), \tag{3}$$

where there are $S \geq 1$ distinct exponential decay terms. A_s describes the initial sound pressure and T_s are decay parameters. The T_s should depend only on the space, and in the case when $S = 1$, they relate to the reverberation time T_{60} as $T_{60} = 13.8T_1^{-1}$. While this is an established model for the sound pressure decay, other models may be developed and can be used interchangeably. The model as described in the rest of the paper is not influenced by what model is chosen for the deterministic decay.

We now turn to the noise terms e^2. The terms $e(t_j)$ are already assumed to be zero-mean, but we additionally assume that they are Gaussian distributed (according to the Central limit theorem) and uncorrelated, with constant variance σ^2. From these assumptions, we can form terms $\varepsilon_j \in \mathcal{N}(0,1)$ by setting

$$\begin{aligned} \varepsilon_j = \frac{e(t_j)}{\sigma} &\Leftrightarrow e^2(t_j) = \sigma^2 \varepsilon^2 \\ \Rightarrow d(t_k) &= \sum_{s=1}^{S} A_s \exp\left[-T_s t_k\right] + \sigma^2 \sum_{j=k}^{K} \varepsilon_j^2, \end{aligned} \tag{4}$$

where $\sum_{j=k}^{K} \varepsilon_j^2 \in \chi^2(K-k+1)$.

This already establishes a value for the likelihood function for the parameters $\theta = \left[\sigma^2, A_s, T_s\right]$ for each value of the Schroeder decay curve. From this point, we form a sequence of prediction residuals r_k which will form the realisation of a stochastic process. The aggregate likelihood for the residuals will form the likelihood function for the parameters θ for the Schroeder curve.

Using the model presented in Eqn. (4), we construct the prediction

$$\hat{d}(t_k) = \mathbb{E}\{d(t_k)\} = \mathbb{E}\{\sum_{s=1}^{S} A_s \exp\left[-T_s t_k\right] + \sigma^2 \sum_{j=k}^{K} \varepsilon_j^2\} = \sum_{s=1}^{S} A_s \exp\left[-T_s t_k\right] + \sigma^2 (K-k+1). \tag{5}$$

This prediction is very similar to the model proposed by Xiang[4], and coincides for certain values of $\Delta t = t_{k+1} - t_k, K$ and σ^2.

The prediction residual for each time point is then calculated as

$$r_k = d(t_k) - \hat{d}(t_k) = \sigma^2 \left(\sum_{j=k}^{K} \varepsilon_j^2 - (K-k+1) \right), \tag{6}$$

and the cumulative distribution function for r_k is given as

$$\begin{aligned} P(r_k \leq y_k) = P\left(\sigma^2 \left(\sum_{j=k}^{K} \varepsilon_j^2 - K + k - 1\right) \leq y_k\right)\Bigg|_{\sigma^2>0} P &= \left(\sum_{j=k}^{K} \varepsilon_j^2 \leq \frac{y_k}{\sigma^2} + K + 1 - k\right) \\ \Leftrightarrow f_{r_k}(y_k) &= \frac{1}{\sigma^2} f_{\chi^2_{n_k}} \left(\frac{y_k}{\sigma^2} + K + 1 - k\right), \end{aligned} \tag{7}$$

where we introduce the notation $f_{\chi^2_{n_k}}$ as the probability density function for a χ^2-distributed variable with $n_k = K - k + 1$ degrees of freedom. This equation will hold for each prediction residual sample r_k independently. In order to determine the joint probability density function, we consider the dependence between two consecutive samples of the residual, r_k, r_{k+1}.

$$r_k = \sigma^2 \left(\sum_{j=k}^{K} \varepsilon_j^2 - K - 1 + k \right) = \sigma^2 \left(\varepsilon_k^2 + \sum_{j=k+1}^{K} \varepsilon_j^2 - K - 1 + k \right) = \sigma^2 \left(\varepsilon_k^2 - 1 \right) + r_{k+1}. \quad (8)$$

Of course, a similar operation could be performed for any $r_k, r_{k+l}, l \leq K - k$. However, when the value r_{k+1} is known, this operation would not provide any additional information. In this sense, the (backwards) sequence $\{r_k\}_{k=K}^{0}$ is Markov, and $f_{r_k|r_K,\ldots,r_k+1} = f_{r_k|r_{k+1}}$. This is due to the independence between noise terms $e(t_k), e(t_j)$ for $k \neq j$.

We now move to determine the conditional probability density function $f_{r_k|r_{k+1}} = f_{r_k|r_{k+1},\ldots r_K}$ where the equality holds by the Markov property.

$$\begin{aligned} P\left(r_k \leq y_k \middle| r_{k+1} = y_{k+1}\right) &= P\left(\varepsilon_k^2 \leq \frac{1}{\sigma^2}\left(y_k - y_{k+1}\right) + 1\right) \\ \Rightarrow f_{r_k|r_{k+1}}\left(y_{k-1}\right) &= \frac{1}{\sigma^2} f_{\chi_1^2}\left(\frac{1}{\sigma^2}\left(y_k - y_{k+1}\right) + 1\right), \end{aligned} \quad (9)$$

with the same notation $f_{\chi^2_{n_k}}$ as before.

Using the Markov property and the definition of conditional probability, the joint probability density function for the sequence $\{r_k\}$ is given as

$$L(\theta) \propto f_{r_0,r_1,\ldots,r_K} = f_{r_0|r_1} \cdot f_{r_1,\ldots,r_K} = \prod_{k=0}^{K-1} f_{k|k+1} \cdot f_{r_K}, \quad (10)$$

where Eqns (7) and (9) give

$$\begin{aligned} L(\theta) \propto \prod_{k=0}^{K-1} f_{r_k|r_{k+1}} \cdot f_{r_K} &= \left(\frac{1}{\sigma^2}\right)^{K+1} \prod_{k=0}^{K-1} f_{\chi_1^2}\left(\frac{1}{\sigma^2}\left(y_k - y_{k+1}\right) + 1\right) \cdot f_{\chi^2_{n_K}}\left(\frac{y_K}{\sigma^2} + 1\right) \\ &= \Big|_{n_K=1} \left(\frac{1}{\sigma^2}\right)^{K+1} \prod_{k=0}^{K-1} f_{\chi_1^2}\left(\frac{1}{\sigma^2}\left(y_k - y_{k+1}\right) + 1\right) \cdot f_{\chi_1^2}\left(\frac{y_K}{\sigma^2} + 1\right). \end{aligned} \quad (11)$$

This function can be iteratively calculated for each set of parameters θ by evaluating the function at the residuals $r_k = d(t_k) - \hat{d}(t_k)$.

As defined, the set of parameters θ over which we evaluate the likelihood function include the variance of the background noise σ^2, but this parameter is a nuisance parameter with the current exponential decay model. It does not describe the deterministic decay, and so any effort put into optimising it will not provide additional information about the system. As such, we will integrate over it. In order to do so more easily, we will modify the prediction residuals r_k and introduce a variable we will call *exponential estimation error*:

$$r_k = d(t_k) - \hat{d}(t_k) = d(t_k) - \left(\sum_{s=1}^{S} A_s \exp\left[-T_s t_k\right] + \sigma^2 \left(K - k + 1\right) \right) = z_k - \sigma^2 \left(K - k + 1\right), \quad (12)$$

where we define $z_k = d(t_k) - \sum_{s=1}^{S} A_s \exp\left[-T_s t_k\right]$ as the exponential estimation error, which does not depend on the parameter σ^2. When this is inserted into Eqn. (11), it becomes

$$L(\theta) \propto \left(\frac{1}{\sigma^2}\right)^{K+1} \prod_{k=0}^{K-1} f_{\chi_1^2}\left(\frac{z_k - z_{k+1}}{\sigma^2}\right) f_{\chi_1^2}\left(\frac{z_K}{\sigma^2}\right). \tag{13}$$

From this equation, we form the log-likelihood. As the objective of this function is optimization over a set of parameters, constants will be discarded. The definition for the χ^2 distribution yields:

$$l(\theta) = -(K+1)\ln\sigma^2 + \sum_{k=0}^{K-1}\left(-\frac{1}{2}\left(\ln(z_k - z_{k+1}) - \ln\sigma^2\right) - \frac{z_k - z_{k+1}}{2\sigma^2}\right) - \frac{1}{2}\left(\ln z_K - \sigma^2\right) - \frac{z_K}{2\sigma^2}. \tag{14}$$

As previously mentioned, σ^2 is a nuisance parameter and should be integrated over. We assign it a uniform prior in the interval $[a, b]$, and obtain the log-likelihood for the interesting parameter set $\phi = \{A_s, T_s\}$:

$$l(\phi) = \frac{z_0}{2}\left(\ln b - \ln a\right) + \frac{\ln z_K}{2}\left(a - b\right) + \frac{a-b}{2}\sum_{k=0}^{k-1} \ln\left(z_k - z_{k+1}\right). \tag{15}$$

With this expression, the likelihood for a set of parameters depends on the difference between the exponential estimate and the measurements at the start and the end of the measurement series, but for the data points in the interior only the difference between two consecutive exponential estimation errors z_k are relevant. These can be interpreted as something of an estimate of the noise terms, and it is reasonable to require these, random data points, to follow a given distribution.

We also consider the influence of the parameters a, b. These can be interpreted as something of a set of weighting parameters. For the second two terms (note that $a < b$), a large difference between these two imply that a large importance is put into having small differences between sequential exponential errors and the final measurement point being very close to the exponential estimate. In the first term, the difference between the logarithms is instead the relevant factor. Here, a large difference implies that the first exponential estimation error should be large. As the differences in the logarithmic domain and in the linear domain will not follow the same pattern, these parameters can be adapted to fit with the data. If the background noise is more or less known, the parameters can be chosen to coincide with, for example, the 95% confidence interval for the background noise level.

2.3 A method to deal with inconsistent data

For measurements with low levels of background noise, or a sufficiently fast decay, this method works well. In other cases, some issues can occur when the difference between two consecutive exponential estimation errors is negative, corresponding to the lower bound defined for the prediction residuals being violated. In these cases, Eqn. (15) will evaluate the logarithm of a negative value, and the result will be undefined. This will always occur for poor choices of parameters, but there exist data sets where no parameter values will produce a valid likelihood. This will occur if the data set contains two consecutive samples which are equal, $d(t_k) = d(t_{k+1})$.

In order to calculate a probability for the measurement series even when this occurs, it is proposed that some measurement points is removed.

For a given set of measurements of a Schroeder decay curve, $\{d(t_k)\}_{k=0}^{K}$, divide the data into two sets M and N, such that

$$\begin{aligned} M &= \{d(t_m) : d(t_{m-1}) - d(t_m) > 0 \cup d(t_0)\} \\ N &= \{d(t_n) : d(t_{n-1}) - d(t_n) \leq 0\}. \end{aligned} \tag{16}$$

We will then discard all measurements from the set N. This will create gaps in the data set, where some points $d(t_m)$ exist but their preceding points $d(t_{m-1})$ do not. For each $m, d(t_m) \in M$ we define $\alpha_m \in \mathbb{Z}_+$, such that α_m is the smallest value such that $d(t_m - \alpha) \in M$. In most cases, α_m will be equal to 1, and no data points will have been discarded. With this definition, it is possible that a sequence at the end of the measurement series is discarded, and in future notation we will use K to denote the last measurement point in M.

Based on the earlier remark that the probability function for each data point could be constructed as being conditional on any of the later time points, we bridge the various gaps of length α. Initially, we note that

$$r_k = \sigma^2 \left(\sum_{j=k}^{K} \varepsilon_j^2 - (K - k + 1) \right) = r_{k+l} + \sigma^2 \left(\sum_{j=k}^{k+l-1} \varepsilon_j^2 - l \right). \tag{17}$$

If we consider the conditional distribution function, we get

$$\begin{aligned} P\left(r_k \leq y_k | r_{k+l} = y_{k+l}\right) &= P\left(\sum_{j=k}^{k+l-1} \varepsilon_j^2 \leq \frac{y_k - y_{k+1}}{\sigma^2} + l \right) \\ \Rightarrow f_{r_k | r_{k+l}}(y_k) &= f_{\chi_l^2}\left(\frac{y_k - y_{k+l}}{\sigma^2} + l \right) \frac{1}{\sigma^2}. \end{aligned} \tag{18}$$

Using the same procedure as with the full data set, we obtain the log-likelihood for the data set M. Note that the partitioning was performed based only on the data set and does not depend on the parameters ϕ.

$$l(\phi) = \frac{z_0}{2}(\ln b - \ln a) + \frac{\ln z_K}{1}(a - b) + (b - a) \sum_{m, \alpha_m} \left(\frac{\alpha_m}{2} - 1 \right) \ln\left(z_m - z_{m+\alpha_m}\right). \tag{19}$$

3 MEASUREMENTS

The methods presented above was applied to measurement data gathered by the author in November 2017. The location for the measurements was the Vadstena Abbey Church in southern Sweden.

While the church was finished in the early 15th century, parts of it dates back to the 14th century. Its stone walls measure about 16m high, and it has a floor area of about $55 \times 33\text{m}^2$.

Reverberation time measurements were performed using both the interrupted noise and the impulse response method, in a total of 7 source positions and 10 receiver positions. The interrupted noise measurements were only considered as control data in this project and are not further discussed. Of the 70 data points, 4 were discarded due to severe inconsistencies. The measurements were performed using a Bruel&Kjaer hand-held analyzer 2270, and paper bags were used as a source of an impulsive impact sound. Data was collected in third-octave bands for the frequency range 50Hz$-$5kHz.

During the time for the measurements, the church was open for visitors and some maintenance work was performed in the church yard. These factors lead to a significant amount of background noise.

4 IMPLEMENTATION

Functions for calculating the log-likelihood were implemented in `MATLAB`, and its dependence on the various parameters and variables were examined graphically, the results of which are discussed in section 5.1. Initial values for the decay parameters were chosen based on estimates of the reverberation time by the classical method.

For parameter and data set combinations where the log-likelihood was not defined (when some of the differences between consecutive exponential estimation errors are negative) the likelihood was disregarded.

5 RESULTS AND ANALYSIS

The (single-valued) reverberation time as estimated with the different methods is presented below.

Values not marked as obtained using the method described in this article have been produced using the Bruel&Kjaer software *Qualifier Type 7839*[c].

5.1 The loglikelihood function surface

The initial step was to further examine the loglikelihood function formed for this particular model, and study how it depends on the parameter values ϕ for a given data set. In order to graphically present the function, only a single slope decay model was considered. The region for which the log-likelihood is not defined is shown as having a zero log-likelihood value.

The log-likelihood function is continuous over the definition set, and there appears to exist an interior maximum close to the values of the parameter T suggested by classical methods. However, there are clear differences between the likelihood values for the different measurement configurations, shown both in figs. 1 and 2. This could be the indication of a significant difference in the sound decay behavior between various positions in the room. Furthermore, the presence of multiple peaks in the log-likelihood function for configuration 20 could indicate that the method is suitable for estimating multi-slope decay patterns.

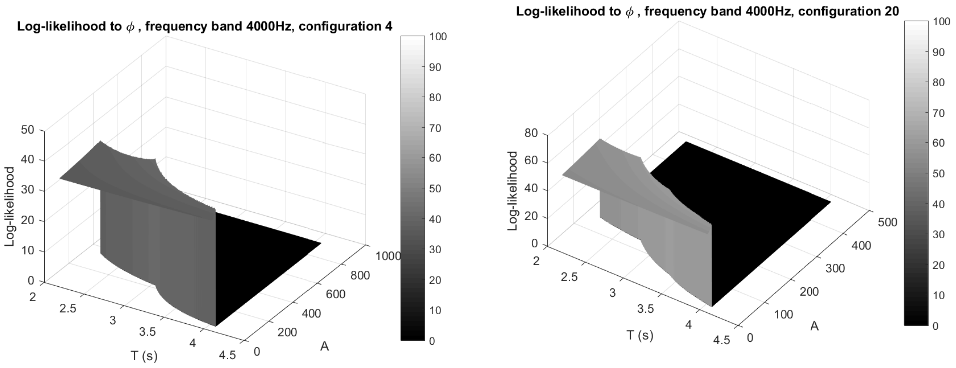
Figures 1 and 2 also give some insight to how the parameters ϕ interact. It seems that for any given A, the likelihood is maximized by choosing the maximum T for which the log-likelihood is defined, and vice versa.

5.2 Exponential sound decay parameter estimates

Some estimates of the exponential sound decay parameters were constructed based on the new model and compared to results from classical methods.

---

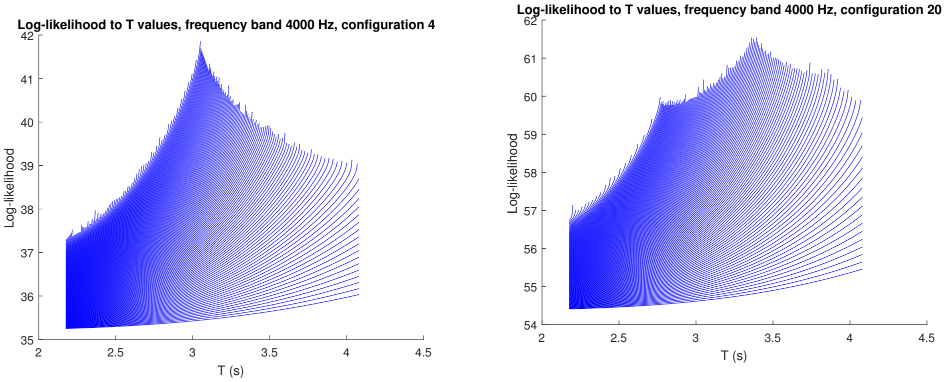
[c] The manual can be found at https://www.bksv.com/downloads/7830/user-manuals/7830-manual-english.pdf



1.1 – The log-likelihood surface as a function of parameter values $\phi = \{A_1, T_1\}$ for the frequency band around $4kHz$ and one of the configurations.

1.2 – The log-likelihood function surface as a function of parameter values $\phi = \{A_1, T_1\}$, for the frequency band around $4kHz$ and one of the configurations.

Fig. 1 – Some examples of the log-likelihood function surface, for one of the frequencies. The function values have been evaluated over a grid of parameter values A_1, T_1. As the measurement series are not identical, the log-likelihood function values are not identical, but some general remarks can be made. The likelihood value increases as A increases, up until the threshold where the likelihood becomes undefined. Furthermore, there seems to exist one or more peaks in likelihood as T varies, showing that the method can be used to produce maximum-likelihood estimates of the reverberation time.



2.1 – Log-likelihood values projected onto the T-axis. Each line represents a different value of A. There is is clear peak for around 3s

2.2 – Log-likelihood values projected onto the T-axis. Each line represents a different value of A. There are a few values T for which there seems to be a peak in the log-likelihood.

Fig. 2 – Graphs showing the log-likelihood projected onto the T-axis. There are clear differences between the graphs, despite them illustrating the behaviour for the same frequency.

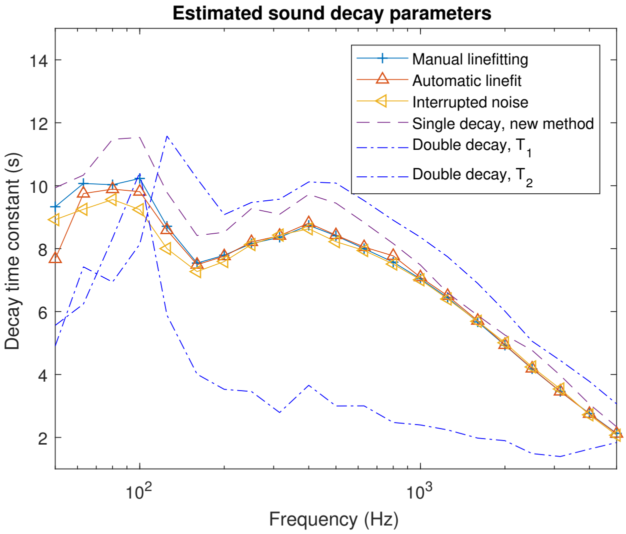


Fig. 3 – Graph showing sound decay parameters for third-octave bands, as estimated by a few different methods. For all cases except the multislope decay, this parameter coincides with the estimated reverberation time. The results in the low frequencies are inconsistent between methods, but the results for the single-slope in the new method over-estimate the reverberation time compared to the classical methods. The double slope decay parameters seem significantly different.

The estimates were constructed by evaluating the log-likelihood function over a equidistant grid over the parameters ϕ for each measurement series. The mean value of the parameters $\phi_{\max}$ was taken as the estimate of the sound decay parameters for this measurement series.

In the case of multi-slope decay scenarios, the ratio between the A_s was used as a parameter, rather than the absolute value. This is consistent with the assumptions made in the introduction, regarding the energy distribution in the case of multi-slope decays.

The estimates for the sound decay parameters T are presented in fig. 3. The method presented in this paper tends to produce a value larger than the other methods. Due to the interaction between the parameters presented in section 5.1, it is possible that this is caused by poor initialisation of the parameter A, as being too small.

The parameter estimates for the curved decay model are significantly different, indicating that they are likely caused by some structure in the data rather than numerical variations. However, the values in the low frequencies seem too low to be realistic. This could, again, be due to poor values of the A parameters.

6 DISCUSSION

The model presented in this paper provides a way to calculate the log-likelihood for any set of parameters describing a deterministic decay, based on measurements of the Schroeder decay curve. The method can easily be adapted to any sound decay model, as it relies on the distribution of the background noise rather than the properties of the sound decay. In the work presented in this paper, the method was tested on exponential sound decays with one or two terms. It was found that the parameters used for the exponential sound decay interact in a way that makes the initialization of the parameters important, and a good and stable optimization algorithm should be used in order to find the maximum.

In the single slope-case, the A-parameter can be considered a nuisance parameter, as it relates to the power of the sound source and the background noise level, rather than any properties of the room. As such it would be reasonable to integrate over it, and it is likely to provide a more numerically stable function. However, the integration limits in this case should depend on the definition set of the log-likelihood function which is very complex.

It is possible that the problem would be solved by just finding the definition boundary, based on how the probability surfaces look. As the likelihood seems to increase as the parameters A grows within the definition surface, finding the value T which allows for the maximum value A could provide the maximum estimate.

In conclusion, the method seems useful for different models of sound decay. However, as the boundaries of the definition set of the log-likelihood function is in general very complicated, care should be put into the optimization algorithms used. Additionally, the interaction between the parameters should be considered before applying this method, and in some cases the results may depend on the initial values to a significant extent.

References

[1] H. Kuttruff. *Room Acoustics, Sixth Edition*. CRC Press, Boca Raton, FL, 2016. ISBN 9781482260434.

[2] M. Vorländer. *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. RWTHedition. Springer Berlin Heidelberg, 2007. ISBN 9783540488309.

[3] N. Xiang. Evaluation of reverberation times using a nonlinear regression approach. *The Journal of the Acoustical Society of America*, 98(4):2112–2121, 1995. doi: 10.1121/1.414460.

[4] N. Xiang and P. M. Goggans. Evaluation of decay times in coupled spaces: Bayesian parameter estimation. *The Journal of the Acoustical Society of America*, 110(3):1415–1424, 2001. doi: 10.1121/1.1390334.