Finite-difference time-domain method for room acoustic simulation - auralization and applications

Jukka Saarelma
Finite-difference time-domain method for room acoustic simulation - auralization and applications

Jukka Saarelma

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Science, at a public examination held at the lecture hall TU2 of the school on the 16th of August 2019, at 12 noon.

Aalto University
School of Science
Department of Computer Science
Virtual Acoustics
Supervising professor
Professor Lauri Savioja

Thesis advisors
Dr. Jonathan Botts
Dr. Pierre Chobeau

Preliminary examiners
Dr. Maarten van Walstijn, Queen's University Belfast, United Kingdom
Associate Professor Cheol-Ho Jeong, Technical University of Denmark, Denmark

Opponent
Professor Michael Vorländer, RWTH Aachen University, Germany

Aalto University publication series
DOCTORAL DISSERTATIONS 137/2019

© 2019 Jukka Saarelma

ISBN 978-952-60-8657-6 (pdf)
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Unigrafia Oy
Helsinki 2019

Finland
Author
Jukka Saarelma

Name of the doctoral dissertation
Finite-difference time-domain method for room acoustic simulation - auralization and applications

Publisher School of Science

Unit Department of Computer Science

Series Aalto University publication series DOCTORAL DISSERTATIONS 137/2019

Field of research Acoustics

Permission for public defence granted (date) 26 June 2019

Language English

Monograph Article dissertation Essay dissertation

Abstract
Room acoustic simulation have been a practical tool in the design of acoustic spaces for several decades. More recently common prediction methods have been optimized and increasingly applied to real-time systems such as virtual reality and augmented reality applications. Modern computing hardware has made approaches that were previously too computationally intensive applicable for wide bandwidth simulation.

Wave-based methods, and especially time-domain solvers for scalar wave equation have gained increasing attention for the use in room acoustic simulation. One of the methods, the finite-difference time-domain method has been proposed as a potential candidate due to convenient scalability properties and the parallel nature of the explicit forms of the method. However, the finite-difference time-domain method has several problematic properties that limit its usage. Due to its direct discretization approach, the dispersion relation in the numerical domain is anisotropic and nonlinear, thus resulting in dispersive wave propagation. Although the properties of the method have been well studied, the implications for its use in auralization have received limited attention.

This thesis studies the numerical error involved in the finite-difference time-domain method within the context of auralization. The absolute thresholds of dispersion error are measured for several different explicit finite-difference schemes as well as under different acoustic conditions. Additionally the method is studied in terms of its application for real-time auralization and analysis of modal time evolution in room acoustics.

The results indicate that the audibility of the dispersion error is dependent on simulation distance, which determines the amount of the accumulated group delay error in the simulation result. The group delay error is masked inaudible by the absorption of air when the phase velocity error is less than 0.28%. Early reflections are shown to further mask the dispersion error. Usage of the method for real-time virtual reality applications with a distributed approach is found to be limited by the synchronization latencies of the computing hardware. The proposed modal time-evolution analysis offers a new approach for evaluating the time-domain characteristics of local interference patterns in general room geometries that can be considered very difficult with existing approaches.

Keywords Finite-difference time-domain method, auralization, room acoustics


ISSN (printed) 1799-4934 ISSN (pdf) 1799-4942

Location of publisher Helsinki Location of printing Helsinki Year 2019

Tekijä
Jukka Saarelma

Väittöskirjan nimi
Aikatason differenssimenetelmä huoneakustiikan simuloinnissa - auralisaatio ja sovellukset

Julkaisija
Perustieteiden korkeakoulu

Yksikkö
Tietotekniikan laitos

Sarja
Aalto University publication series DOCTORAL DISSERTATIONS 137/2019

Tutkimusala
Akustiikka

Käsikirjoituksen pvm
08.05.2019

Väittöspäivä
16.08.2019

Väittelyluvan myöntämispäivä
26.06.2019

Kieli
Englanti

Monografia
Artikkeliväittöskirja

Avainsanat
Aikatason differenssimenetelmä, auralisaatio, huoneakustiikka

ISBN (painettu)
978-952-60-8656-9

ISBN (pdf)
978-952-60-8657-6

ISSN (painettu)
1799-4934

ISSN (pdf)
1799-4942

Julkaisupaikka
Helsinki

Painopaikka
Helsinki

Vuosi
2019

Sivumäärä
129

urn
Preface

The research work for the results that are presented in this thesis has been carried out at the Department of Computer Science and the former Department of Media Technology at the Aalto University School of Science in Finland. The work was funded by the Academy of Finland (project no. 265824). Part of the research was conducted during an internship at Oculus Research in 2017.

First and foremost I would like to thank my thesis supervisor Prof. Lauri Savioja for continuous encouragement, support and advice. It has been a privilege to work in such a free, yet supportive environment. I would like to thank my thesis advisors Dr. Jonathan Botts and Dr. Pierre Chobeau for all the support during the different stages of the research work.

I wish to thank the pre-examiners of this thesis, Dr. Maarten van Walstijn and Prof. Cheol-Ho Jeong, for their feedback on the manuscript.

I am grateful to all former and current members of the virtual acoustics research group at Aalto University for creating an inspiring work atmosphere. Prof. Tapio Lokki, Dr. Sakari Tervo, Dr. Jukka Päätynen, Dr. Antti Kuusinen, Dr. Aki Haapaniemi, Henna Tahanainen, Dr. Robert Albrecht, Kai Saksela, Dr. Philip Robinson, Dr. Samuel Siltanen, Dr. Alex Southern, Perttu Laukkanen, Otto Puomio, Julie Meyer and Raimundo Gonzalez are all thanked for the countless discussions that have given perspective and motivation, and for showing by example how research can be done. Special thanks go to my colleague Sebastian Prepeliţă for all the discussions, and namely full-on rants on various aspects of research. They have most definitely raised the bar for quality in all aspects of the work.

I would like to thank Dr. Ravish Mehra for various insights regarding product oriented research, and Jonathan Califa for professionalism and excellent morale during the long days of the proof-of-concept development. I also wish to thank Dr. Brian Hamilton for the exemplary attention to detail, and the numerous discussions on room acoustic simulation. All the rants were much appreciated.
Preface

My sincere thanks go to the members of the former pub quiz team Karhucopla, various fellow climbers, and all friends for much needed distraction at times when the work may have felt less inspiring. Finally, I wish to thank my parents for being firm proponents of education at all stages of my life. It may have been worthwhile.

Helsinki, July 1, 2019,

Jukka Saarelma
Preface 1
Contents 3
List of Publications 5
Author’s Contribution 7
1. Introduction 13
  1.1 Scope of the Thesis 15
  1.2 Organization of the Thesis 16
2. Finite-difference time-domain method 17
  2.1 Relation to other room acoustic simulation methods 18
    2.1.1 Geometric acoustic methods 18
    2.1.2 Wave-based methods 19
  2.2 Propagation model 19
    2.2.1 Sound propagation in air 19
    2.2.2 Sound incidence at a boundary 20
  2.3 Compact explicit finite-difference time-domain schemes 21
    2.3.1 Boundary conditions 23
  2.4 Numerical dispersion 28
  2.5 Computational considerations 31
3. Experimental methods for perceptual evaluation of room acoustic simulation 35
  3.1 Auralization methods 36
    3.1.1 Binaural technique 37
    3.1.2 Loudspeaker reproduction 39
    3.1.3 Spatial encoding 40
  3.2 Threshold measurement 43
    3.2.1 Psychometric function 43
    3.2.2 Adaptive procedures 44
This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.


Author’s Contribution

Publication I: “Audibility of dispersion error in room acoustic finite-difference time-domain simulation as a function of simulation distance”

The threshold of the dispersion error was measured for finite-difference time-domain (FDTD) simulation in free field conditions. The study considers three different compact explicit FDTD schemes. The level of dispersion is controlled by changing the theoretical simulation distance, which is evaluated using a dispersion filter for representing the error introduced to a propagating plane-wave inside a simulation domain in the worst-case direction of each scheme. The measured thresholds indicate that the maximum group-delay value associated with the error is similar to the thresholds for group-delay previously reported in the literature, and that the error behaves similarly in all measured schemes.

The present author designed the experiment, implemented all code concerning stimuli generation, conducted the listening tests and wrote 90% of the manuscript.

Publication II: “Audibility of dispersion error in room acoustic finite-difference time-domain simulation in the presence of absorption of air”

An obvious question that remained open after Publication I concerned the effect of air absorption on the audibility of dispersion error, as the dispersion error in the studied compact explicit schemes increases as a function of frequency. The experiment was designed to measure the audibility of dispersion error in free-field when absorption of air is added to the signal. Only one compact explicit FDTD scheme was used in the experiment as the first experiment indicated that the error is similar for all three schemes of interest. The amount of error was controlled by adjusting the phase velocity error percentage, and the threshold
was measured at four different distances with absorption of air included and excluded. The results indicated that the error becomes inaudible when the phase velocity error percentage falls below 0.28%.

The present author designed the experiment, implemented all code concerning stimuli generation, conducted the listening tests and wrote the manuscript.

**Publication III: “Challenges of Distributed Real-Time Finite-Difference Time-Domain Room Acoustic Simulation for Auralization”**

There has been a growing interest in real-time wave-based room acoustic simulation for applications such as gaming and virtual reality (VR). Even though the FDTD method has been shown to scale well for large problems on distributed hardware, no studies have attempted to accelerate the method up to real-time constraints using smaller problem domains and distributed general purpose graphics processing unit devices. In order to study the paradigm, a proof-of-concept system for real-time auralization in an VR environment was implemented using distributed FDTD simulation.

The present author implemented the simulation code, implemented the code enabling communication between the distributed system and the VR scene, conducted the experiments evaluating the performance of the system, and wrote the manuscript.

**Publication IV: “Spatial analysis of modal time evolution in room acoustics”**

An attractive property of the FDTD method is that it allows a time-domain solution for every spatial position inside the simulation domain. This property has long been used to visualize the time-domain evolution of sound fields inside various geometric constructs. Although such visualizations may be intuitive for general inspection of sound scattering, they lend little information concerning the effect of interference phenomena at different frequency bands. In this study, an approach was proposed to spatially analyze the propagation of energy inside general geometries in arbitrary frequency bands.

The present author implemented the code related to the analysis method, produced the media content and wrote the manuscript.
After the two first experiments regarding the audibility of dispersion error, some questions that remained were: how do early reflections affect the audibility of the dispersion error? and how do the measured thresholds predict the audibility of the error in auralized, simulated room responses? Two experiments were conducted to address these questions. First experiment measured the audibility of the dispersion error in a case where the subject is presented a response containing the direct sound and a single early reflection. Thresholds were measured for the cases \textit{dispersion error in direct sound} and \textit{dispersion error in the first reflection} in different conditions. The second experiment measured the audibility of the dispersion error in an auralized simulated room response. The measured thresholds indicated that the early reflection does increase the threshold in comparison to free field conditions. The results of the second experiments were inconclusive.

The present author designed the experiment, implemented all code concerning stimuli generation, conducted the listening tests and wrote the manuscript.
List of abbreviations

3-D  Three-dimensional
AR  Augmented reality
ARD  Adaptive rectangular decomposition
BEM  Boundary element method
BRIR  Binaural room impulse response
BTL  Brandley-Terry-Luce
CCP  Close cubic-packed
DirAC  Directional audio coding
DOA  Direction of arrival
DWM  Digital waveguide mesh
ESM  Equivalent source method
FCC  Face centered cubic
FDTD  Finite-difference time-domain
FEM  Finite element method
FVTD  Finite-volume time-domain
GA  Geometric acoustics
HRTF  Head related transfer function
IWB  Interpolated wide band
List of abbreviations

**JND**  just noticeable difference

**PDF**  probability density function

**PSTD**  Pseudo spectral time-domain

**SDM**  Spatial decomposition method

**SH**  Spherical harmonic

**SHT**  Spherical harmonic transform

**SIRR**  Spatial impulse response rendering

**SRL**  Standard rectilinear

**VR**  Virtual reality

**WFS**  Wave field synthesis
1. Introduction

Room acoustic simulation methods have several applications with various practical aims and requirements. The first example application could be room acoustic prediction in consultant work, where the acoustic properties of different renovation or design options are evaluated. In such applications, physical accuracy of the model is a priority, and computations may be carried out off-line. The second example could be applications where there is no concern about the actual physical interpretation of the model, but an aim to convey a spatial image that evokes the desirable sense of a space with a low computational requirement. Such application could be a reverberation effect that a studio engineer would apply on a track to change the impact of a musician’s performance, or a game engine that would use a simple heuristic approach to represent the reverberation of a space within a game. A third example, which would land somewhere in between the two previous examples, is an application that aims to give as plausible a spatial image as possible of a space with limited computational requirements. Such an application could be a virtual reality (VR) sound engine that produces a representation of a sound field inside a virtual space in real-time, so that the spatial image that is presented to the user would match the expectations of the user as closely as possible, given the presented visual media.

The given examples may be thought of as ranging between two opposite objectives; one with an aim of physical accuracy without strict time constraints, and one aiming at low computing time at the cost of strict physical accuracy while preserving a chosen level of plausibility. For some applications, it may be reasonable to assume that the two ends of the range are converging over time, namely a physically very accurate sound field can be produced within the real-time constraints using future computing hardware and algorithms. However, it should be noted that even if a tool for the creation of physically accurate sound fields in real-time would exist, there should be no expectation that it would be used by a content creator or an artist to create sound fields that represent any real-world conditions. Therefore such convergence may not be considered a universal aim. As noted by Toole about sound reproduction in general[128, p. 6] with citations of the work of Sterne [121]:
Sterne (2003) goes on to explain that “as many critics of film and photography have shown us, reality is as much about aesthetic creation as it is about any other effect when we are talking about media” (p. 241). And, in the context of sound recording, “far from being a reproduction of the actual event, the recording was a ‘re-creation’” (p. 242). The goal is not imitation but the creation of specific listener experiences. This certainly exists dramatically in the directional and spatial experiences in reproduced sounds.

Due to the different aims of the applications and the intentions of content creators, the evaluation of the perceived quality of simulated sound fields can be considered highly subjective and difficult to measure. If an assumption is made that the aim of the acoustic simulation in the application being studied is to acquire a physically accurate representation of the sound field within a space, the evaluation of the perceptual quality may be approached with an increased sense of objectivity. In such a case, the perceptual quality may be quantified to some extent in terms of perceivable error.

The acoustic fields that are a product of numerical simulation can be characterized in a different manner than measured acoustic systems. Firstly, a simulated acoustic sound field is a product of explicitly set initial conditions; the geometry is approximated by the user, the definition of boundary conditions is carried out by the user, the errors related to the used method are often quantifiable, and the source and receiver characteristics are explicitly specified. The numerical solution is assumed to converge towards the chosen theoretical physical model.

On the contrary, real acoustic systems, whether a room, a human head, or a loudspeaker source, may often be decomposed to numerous complex interactions of waves propagating through various heterogeneous mediums, reflecting from boundary layers and then interacting with a device capable of transforming pressure deviations to signals, being human cochlea or a microphone capsule. The real system can only be characterized by making observations of the conditions; measurement of the surface materials, measurement of the atmospheric conditions, knowledge of the electronic parameters of the source, and scattering properties of the receiver. Ideally, the simulation would take into account all possible physical interactions, but commonly simplifications are made [101, pp. 27–28]. Simplifications may be motivated by a small assumed effect on the simulation result, and computational complexity that a more complicated model would entail.

Comparing the perceptual quality of existing acoustic conditions involves data from real physical observations. Therefore the results may be considered meaningful as essentially the comparison is between the physical observations, given that the method of data representation to the listener is the same between the conditions. In the case of acoustic simulation, comparing two numerical solutions of the same method, or possibly the solutions of two different methods, the comparison may involve various different types of errors specific to the methods. These discrepancies may affect the perceived quality either negatively...
or positively in a holistic sense. Therefore, common approaches to assess the simulated solutions are to measure the perceived similarity to real reference condition, or directly quantify a source of error and evaluate its perceptual significance. From these two, equality to a real reference can be considered the definitive goal, but direct similarity measurement without explicit knowledge of the errors involved may lead to lesser generality of the results. Measuring the perceptual significance of a single source of error does not necessarily as such give much information about the plausibility of the final numerical response, but is a solid building block in a generalizable assessment of the method.

With these observations and notions in mind, the approach and the scope of this thesis may be described as follows.

### 1.1 Scope of the Thesis

In this thesis, the focus is placed on the audibility of the numerical error; what, on the perceptual level, are the most problematic errors related to the compact explicit finite-difference time-domain FDTD schemes, and at what level of error is needed for it to go unnoticed. The relation of the error in theoretical models is illustrated in Fig. 1.1. Additionally, studies on the practical use of the FDTD method, and time domain solvers in general are conducted. The studies include the evaluation of the applicability of the FDTD method for distributed real-time auralization systems, and the usage of time-domain solutions in the prediction of the time-evolution of modal components in spaces.

The thesis is limited to compact explicit FDTD schemes due to their scalability properties, and the conducted studies consider only three-dimensional (3-D) numerical solutions.
1.2 Organization of the Thesis

The thesis is organized as follows. The second section reviews the basic principles of compact explicit FDTD schemes, relevant portions of the physics of sound propagation, and the numerical error associated with the studied schemes. In the third section, the experimental methods used to measure different perceptual attributes are reviewed. Finally, in the fourth section, the primary contributions of this thesis are summarized.
2. Finite-difference time-domain method

The FDTD method is a part of a larger family of finite difference methods that are used for solving various partial, and ordinary differential equations. Finite difference methods have been applied to various acoustic wave propagation problems, such as sound synthesis [118, 14, 15, 18, 136], and room acoustic simulation [13]. The basic idea of the FDTD method is to formulate the partial differential equation in terms of difference operators defined in a discretized volumetric grid at discrete points in time. In room acoustic context, difference methods were initially proposed for solving the equations of linear acoustics in the time domain in terms of particle velocity vectors and pressure [22, 23] and since then, different formulations for deriving finite difference schemes have been proposed primarily with respect to scalar variables [105, 104, 86, 17].

The FDTD method has several convenient properties. One of them is that a time-domain solution is directly available in each point of the simulation domain. This allows the analysis and visualization of the complete sound field after a single simulation. Another property that is often used to accelerate the computation in various explicit FDTD schemes is that the simulation update can be done in parallel for each point of the domain. This allows an efficient usage of parallel computing hardware for the simulation [102, 138, 49, 51]. The FDTD method admits the usage of several different boundary condition types [67], and by using an interpretation as the finite-volume time-domain (FVTD) method, compact FDTD schemes can be extended to handle general geometries using fitted boundary cells [16]. Furthermore, it has been shown that FVTD interpretation allows the derivation of stability conditions for various impedance boundary conditions for general geometries [19].

This section reviews the physical models that are used as the basis of the numerical approximation, the basic principles of the used compact explicit FDTD schemes, the quantification of the dispersion error related to these schemes, and finally, computational considerations relevant to the performance of the implementation of the FDTD method in general.
2.1 Relation to other room acoustic simulation methods

The FDTD method is one of many methods that are used for room acoustic prediction problems, which vary in both physical interpretation of the problem and in numerical approximation strategies. In literature, techniques have often been divided into two main categories: geometric acoustics (GA), and wave-based methods. In this section, a brief review of these methods is given.

2.1.1 Geometric acoustic methods

GA methods are computational approaches where the sound propagation is modeled as rays that are traveling in the modeled space from the source position, via various surface reflections to the receiver position. The model is often described as analogous to the propagation of light within a space. The central part of the solution is the path of the casted ray, since it determines how much the sound is attenuated and scattered due to surface interactions. The propagation paths of the rays may be determined with rules of varying levels of complexity. The first proposed computational techniques for GA were ray-tracing [69] and image source method [2, 21]. Both of the early approaches only considered specular reflections, and omitted the diffraction of sound waves. The difference between these two early methods is that ray-tracing is considered stochastic and the image source method deterministic. In ray-tracing, a limited number of rays are casted from the source position, and only a portion of the rays hit the receiver position and hence contribute to the solved response. The ray is interpreted as representing a portion of the total energy included in all casted rays. In the image source method, the paths from the source to the receiver are explicitly solved using the position information of the source and receiver, and the geometric model of the space. In such cases, the solution contains a complete set of image sources up to the given reflection order.

Several more sophisticated GA techniques have been introduced since, for example, methods to include diffracted waves[123, 129, 131], and different approaches to evaluate reflections for rough surfaces [32, 117]. The techniques to optimize the path calculation process has been under increasing interest [40, 131, 42, 73], and other types of optimization regarding efficient evaluation of higher-order diffracted rays [112] and the handing of large numbers of sources using dynamic clustering [111] have been proposed.

GA methods are attractive alternatives for real-time systems [41, 106, 142, 109, 114, 91] due to the fact that simplifications do not strictly limit the frequency bandwidth of the solution. Very simple solutions, such as image source solution with very low reflection order and optimized computing can give approximations covering the whole audible bandwidth and are fast to calculate especially for simplified geometries. This allows a flexible design of algorithms in terms of the available computational resources. The simplifications are made in terms of reflection order and the representation of the late reverberation. The solutions of
GA methods can be generalized with the *room acoustic rendering equation* [117] that underlines the cost of the flexibility; GA methods do not produce numerical solutions to the wave equation other than in very specific cases [103].

### 2.1.2 Wave-based methods

Wave-based methods aim to give a numerical solution to a physical formulation of sound propagation. An often-used division of the methods is in time-domain, and frequency domain solvers. Commonly used frequency domain methods for acoustic problems include the *finite element method* (FEM) [6], and the *boundary element method* (BEM) [52, 56]. Frequency domain solvers use assumptions of time-harmonic signals and following this assumption, use the Helmholtz equation as the basis of the numerical solution. FEM uses volumetric discretization, whereas BEM uses a discretization of the surfaces of the geometry. Time-domain wave-based methods that use a wave number domain representation of the sound field include *adaptive rectangular decomposition* (ARD) [98], *pseudo-spectral time-domain* method (PSTD) [54]. Both of these approaches are time-stepping methods where the spatial domain is divided into rectangular sub-domains in which the sound field can be represented using modal decomposition.

A method closely related to the FDTD method, the FVTD method [22, 16] has been proposed for room acoustic prediction. FVTD differs from FDTD in its basic premise; the volume is not discretized using a discrete grid of points, but as a set of volumes that represent an average of a scalar field within the element. Such an approach leads to flexibility in the definition of boundary conditions as they can be represented by arbitrarily shaped volumes and fluxes between volume boundaries [16] instead of operations between the nodes of a structured grid. The FVTD method also makes it possible to conduct stability analysis for various boundary condition types [19].

A method that may be considered similar to FVTD and FEM is the *discontinuous Galerkin* method [53], that has been proposed for acoustic problems [72, 135]. In simplified terms the main difference between FVTD method and the discontinuous Galerkin method is that, while in FVTD the scalar value within an element is an average represented by a single value, in discontinuous Galerkin, the field variable within an element may be represented with a set of basis functions.

### 2.2 Propagation model

#### 2.2.1 Sound propagation in air

The conservation equations of mass and momentum read:

$$\frac{\partial p}{\partial t} = -\kappa_0 \nabla \cdot \mathbf{v}, \quad (2.1)$$
and

\[ \frac{\partial \mathbf{v}}{\partial t} = -\nabla p, \tag{2.2} \]

where \( \kappa_0 = \rho c^2 \), \( \rho \) is the ambient density, \( p \) is acoustic pressure, \( \mathbf{v} \) particle velocity vector, \( t \) is time, \( c \) the speed of sound, and \( \nabla \) the divergence operator. Now by following Bergmann [10, Section 5], by first dividing Eq. (2.1) by \( \kappa_0 \) and then differentiating with respect to \( t \), we get

\[ \frac{1}{\kappa_0} \frac{\partial^2 p}{\partial t^2} = -\frac{\partial}{\partial t} (\nabla \cdot \mathbf{v}). \tag{2.3} \]

Subsequently dividing Eq. (2.2) with \( \rho \) and taking the divergence operator, we get

\[ \nabla \cdot \left( \frac{1}{\rho} \nabla p \right) = -\nabla \cdot \frac{\partial}{\partial t} \mathbf{v}, \tag{2.4} \]

and further assuming that ambient density \( \rho \) is independent of position,

\[ \frac{1}{\rho} \Delta p = -\nabla \cdot \frac{\partial}{\partial t} \mathbf{v}, \tag{2.5} \]

where \( \Delta \) is the Laplacian. By equating Eq. (2.3) and Eq. (2.5), we arrive at scalar wave equation:

\[ \frac{\partial^2 p}{\partial t^2} = c^2 \Delta p. \tag{2.6} \]

The Laplacian operator \( \Delta \) can be written:

\[ \frac{\partial^2 p}{\partial t^2} = c^2 \left( \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \right). \tag{2.7} \]

The scalar wave equation describes sound propagation in a lossless medium, and it is used as the basis for the numerical solutions in this thesis. Atmospheric absorption is taken into account in a post-processing stage of simulation results, which is described more elaborately in Publication II.

### 2.2.2 Sound incidence at a boundary

When a sound wave propagating in a medium hits a boundary surface, a part of the energy is transmitted or absorbed to the surface structure, and a part of the sound energy is reflected back to the medium. The reflected part of the wave may differ from the incident wave both in amplitude and phase. The reflection characteristics of a surface are commonly described with a reflection coefficient. A reflection coefficient for an infinite plane is defined as:

\[ R(\omega, \theta_i) = \frac{\xi(\omega) \cos(\theta_i) - 1}{\xi(\omega) \cos(\theta_i) + 1}, \tag{2.8} \]
where \( \theta_i \) is the angle of incidence, and
\[
\zeta(\omega) = \frac{Z}{\rho c}, \tag{2.9}
\]
is the specific impedance of the surface, and \( Z \) the surface impedance that determines the relationship of the pressure and the normal component of the particle velocity at the surface. The normal incidence specific surface impedance at the boundary can be expressed as a function of the normal incidence reflection coefficient as
\[
\zeta(\omega) = \frac{1 + R(\omega, 0)}{1 - R(\omega, 0)}. \tag{2.10}
\]
Alternative representation that is commonly used is the absorption coefficient
\[
\alpha = 1 - |R|^2, \tag{2.11}
\]
that indicates the portion of energy that is absorbed by the surface. Both normal incidence absorption and normal incidence reflection coefficients can be expressed in terms of normal incidence surface impedance. For GA applications, it is common to use a random incidence reflection and absorption coefficient, which indicates the average reflection properties of the surface over the incidence angle. This value differs from the normal incidence coefficient that is often used in wave-based methods.

Regarding the used FDTD schemes, an assumption is made that the surfaces are locally reacting. This means that at any surface element, the normal component of the particle velocity does not depend on the pressure at the neighboring surface elements, but only on the pressure value at the element itself [70, p. 44].

### 2.3 Compact explicit finite-difference time-domain schemes

FDTD schemes that use only immediate neighboring cells in the update and use explicit forms of the update are called compact\(^1\) explicit FDTD schemes. Such schemes have had a fair amount of interest in the literature, and a parameterization of groups of such schemes have been proposed [12, 68].

In the following, the standard rectilinear (SRL) scheme is derived as the scheme is used in the simulations of Publication I, IV and V. The two other schemes investigated in Publications I and II are discussed in brief.

A common initial step of finite difference schemes is to represent the domain as a set of discrete points. Here, the interest is a scalar field variable \( p(\mathbf{x}, t) \) at position \( \mathbf{x} \) in a 3-D domain at time \( t \). The continuous variable \( p \) is represented in discrete spatial points \( p^h_{k,l,m} = p([k\Delta X, l\Delta Y, m\Delta Z], n\Delta T) \). A set of second-order

\(^1\)The description compact is used here similarly to Kowalczyk [66, p. 68], but it is noted that the term is not agreed upon in literature as pointed out by Hamilton [48, p. 118]
difference operators $\delta_t^2$, $\delta_x^2$, $\delta_y^2$, and $\delta_z^2$ can now be defined in this discrete domain as follows

\[
\begin{align*}
\frac{\partial^2}{\partial t^2} p_{k,l,m}^n & = \frac{p_{k,l,m}^{n+1} - 2p_{k,l,m}^n + p_{k,l,m}^{n-1}}{\Delta T^2} + O(\Delta T^2), \\
\frac{\partial^2}{\partial x^2} p_{k,l,m}^n & = \frac{p_{k+1,l,m}^n - 2p_{k,l,m}^n + p_{k-1,l,m}^n}{\Delta X^2} + O(\Delta X^2), \\
\frac{\partial^2}{\partial y^2} p_{k,l,m}^n & = \frac{p_{k,l+1,m}^n - 2p_{k,l,m}^n + p_{k,l-1,m}^n}{\Delta Y^2} + O(\Delta Y^2), \\
\frac{\partial^2}{\partial z^2} p_{k,l,m}^n & = \frac{p_{k,l,m+1}^n - 2p_{k,l,m}^n + p_{k,l,m-1}^n}{\Delta Z^2} + O(\Delta Z^2).
\end{align*}
\]  

(2.12)

By hereafter assuming a structured Cartesian grid where $\Delta X = \Delta Y = \Delta Z$, and simply substituting the differential operations with difference operators in Eq. (2.6), a discrete representation of the wave equation is achieved:

\[
\frac{p_{k,l,m}^{n+1} - 2p_{k,l,m}^n + p_{k,l,m}^{n-1}}{\Delta T^2} = c^2 \left( \frac{p_{k+1,l,m}^n - 2p_{k,l,m}^n + p_{k-1,l,m}^n}{\Delta X^2} \right. \\
+ \frac{p_{k,l+1,m}^n - 2p_{k,l,m}^n + p_{k,l-1,m}^n}{\Delta Y^2} \\
+ \left. \frac{p_{k,l,m+1}^n - 2p_{k,l,m}^n + p_{k,l,m-1}^n}{\Delta Z^2} \right),
\]  

(2.13)

from which the future time step $p_{k,l,m}^{n+1}$ can be solved:

\[
p_{k,l,m}^{n+1} = \lambda^2 \left( p_{k+1,l,m}^n + p_{k-1,l,m}^n + p_{k,l+1,m}^n + p_{k,l-1,m}^n + p_{k,l,m+1}^n + p_{k,l,m-1}^n \right) \\
+ (2 - 6\lambda^2) p_{k,l,m}^n - p_{k,l,m}^{n-1},
\]  

(2.14)

where $\lambda = \frac{c\Delta T}{\Delta X}$ is the Courant number. This is the update equation of the SRL scheme in the absence of a boundary.

The two other compact explicit FDTD schemes that were investigated in Publications I and II are close cubic-packed (CCP), (synonymous to face center cubic (FCC)), and interpolated wide band (IWB). Both of the schemes are part of a family of compact explicit schemes [68]. These schemes, similarly to SRL, are second-order accurate, but the difference is in the discrete evaluation of the Laplacian operator. In the SRL scheme, only axial neighbors are used in the approximation of the second-order partial derivatives, whereas in CCP and IWB, a larger set of points is used. The point sets, which are commonly referred to as stencils are illustrated in Figure 2.1.

The update equations for the SRL, CCP and IWB schemes can all be included

22
The parameter values for different compact explicit FDTD schemes.

<table>
<thead>
<tr>
<th></th>
<th>SRL</th>
<th>CCP</th>
<th>IWB</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>1</td>
<td>0</td>
<td>1/4</td>
</tr>
<tr>
<td>b</td>
<td>0</td>
<td>1/4</td>
<td>1/8</td>
</tr>
<tr>
<td>c</td>
<td>0</td>
<td>0</td>
<td>1/16</td>
</tr>
<tr>
<td>d</td>
<td>6</td>
<td>3</td>
<td>7/2</td>
</tr>
<tr>
<td>λ</td>
<td>√1/3</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

2.3.1 Boundary conditions

A boundary can be introduced to a compact explicit FDTD scheme by substituting a node that is just outside the given geometry with a discrete boundary condition. These nodes just outside the discretized geometry are referred to as "ghost
Finite-difference time-domain method

points” [127, p. 15]. The discrete boundary condition for the SRL scheme can be derived by approximating the first-order derivatives in the time and space domains of the continuous boundary conditions with difference operators.

The continuous boundary conditions for different Cartesian coordinate axes in 3-D can be expressed similarly to Eq. (2.21) as

$$p_{n,k,l,m+1} = p_{n,k,l,m} + \frac{1}{\lambda \xi_x} \left( p_{n-1,k,l,m} - p_{n+1,k,l,m} \right), \quad (2.21)$$

where $\xi$ is the specific surface impedance at the boundary.

Continuous boundary conditions can be introduced into the numerical model using a second-order approximation using centered-differences [67]. Centered-difference operators that are used to approximate the first-order derivatives in the boundary conditions Eq. (2.16) read:

$$\frac{\partial p}{\partial t} \approx \delta_t = \frac{p_{n,k,l,m+1}^{n+1} - p_{n,k,l,m+1}^n}{\Delta T} + O(\Delta T^2), \quad (2.17)$$

$$\frac{\partial p}{\partial x} \approx \delta_x = \frac{p_{n,k,l,m+1}^{n+1} - p_{n,k,l,m}^n}{\Delta X} + O(\Delta X^2), \quad (2.18)$$

$$\frac{\partial p}{\partial y} \approx \delta_y = \frac{p_{n,k,l,m+1}^{n+1} - p_{n,k,l,m}^n}{\Delta Y} + O(\Delta Y^2), \quad (2.19)$$

$$\frac{\partial p}{\partial z} \approx \delta_z = \frac{p_{n,k,l,m+1}^{n+1} - p_{n,k,l,m}^n}{\Delta Z} + O(\Delta Z^2). \quad (2.20)$$

The ghost point in the direction of the negative x-axis of a space may be expressed using the difference operator in the x-dimension:

$$p_{n,k-1,l,m} = p_{n,k+1,l,m} + \frac{1}{\lambda \xi_x} \left( p_{n-1,k,l,m}^{n} - p_{n+1,k,l,m}^{n} \right), \quad (2.21)$$

The solved node can be substituted in the air update Eq. (2.14), from which an update equation for the boundary node as illustrated in Figure 2.2 as x-boundary is formulated as

$$p_{n,k,l,m}^{n+1} = \left[ \lambda^2 \left( p_{n,k+1,l,m}^{n} + p_{n,k-1,l,m}^{n} + p_{n,k+1,l,m+1}^{n} + p_{n,k+1,l,m-1}^{n} + 2p_{n,k+1,l,m}^{n} \right) \right] + (2 - 6\lambda^2) p_{n,k,l,m}^{n} + \left( \frac{\lambda}{\xi_x} - 1 \right) p_{n-1,k,l,m}^{n} \left( 1 + \frac{\lambda}{\xi_x} \right). \quad (2.22)$$

In the case of a corner node, boundary conditions in the direction of both, the x- and y-axis, have to be met simultaneously. With a procedure similar to the previously used, the ghost points in the direction of x- and y-axis are substituted with the corresponding boundary conditions.

The equation for the ghost point in the direction of positive y-axis can be expressed similarly to Eq. (2.21) as

$$p_{n,k,l+1,m} = p_{n,k,l-1,m} + \frac{1}{\lambda \xi_y} \left( p_{n-1,k,l,m}^{n} - p_{n+1,k,l,m}^{n} \right), \quad (2.23)$$
where $\xi_y$ is the specific acoustic impedance of the boundary towards the positive y-axis. With substitutions to Eq. (2.14), an update for the condition for a node with boundary conditions in the direction of x- and y-axis as illustrated in Figure 2.2 as x-y-corner is given

$$
p_{k,l,m}^{n+1} = \left[ \frac{\lambda^2 (2p_{k+1,l,m}^n + 2p_{k,l-1,m}^n + p_{k,l,m+1}^n + p_{k,l,m-1}^n) + (2 - 6\lambda^2)p_{k,l,m}^n}{1 + \frac{\lambda}{\xi_x} + \frac{\lambda}{\xi_y}} \right].
$$

(2.24)

A similar approach can be applied to a case where the node contains boundaries in the direction of each Cartesian coordinate axes. According to this boundary scheme, re-entrant corners do not have ghost points to eliminate, and therefore the node should be updated with the update equation for air node. However, as can be observed from Figure 2.2, the geometric interpretation of this boundary model does not directly agree with such convention. As pointed out by Hamilton [47], this centered-difference boundary model may become unstable in certain geometric configurations, and proposes an approach to derive updates for re-entrant corners using a finite-volume interpretation. By using this approach, an example update for a re-entrant corner where one-quarter node locates in the direction of positive y-axis and negative x-axis (as in Figure 2.2) can be written:

$$
p_{k,l,m}^{n+1} = \left[ \frac{2\lambda^2 (2p_{k+1,l,m}^n + p_{k,l-1,m}^n + p_{k,l,m+1}^n + 2p_{k,l,m}^n) + \left( \frac{\lambda}{\xi_x} + \frac{\lambda}{\xi_y} - 1 \right)p_{k,l,m}^{n-1}}{1 + \frac{2\lambda}{3\xi_x} + \frac{2\lambda}{3\xi_y} - 1} \right].
$$

(2.25)

Similar updates for re-entrant corners for different Cartesian coordinate axes may be derived using the approach. This boundary model is not considered in the studies of this thesis.

An alternative formulation for the boundary that leads to a convenient update equation for parallel computing can be derived by approximating the continuous boundary condition 2.16 with a first-order difference operator in space instead of the centered-difference operator. This boundary type has also been referred to as velocity-centered boundary condition and it has been shown to alleviate potentially unstable modes in room simulation [24]. First-order difference
operators that are considered here in the 3-D grid are defined as:

\[
\begin{align*}
\frac{\partial p}{\partial x} & \approx \delta_{x+} = \frac{p_{k+1,l,m}^n - p_{k,l,m}^n}{\Delta X} + O(\Delta X) \\
\frac{\partial p}{\partial y} & \approx \delta_{y+} = \frac{p_{k,l+1,m}^n - p_{k,l,m}^n}{\Delta X} + O(\Delta X) \\
\frac{\partial p}{\partial z} & \approx \delta_{z+} = \frac{p_{k,l,m+1}^n - p_{k,l,m}^n}{\Delta X} + O(\Delta X) \\
\frac{\partial p}{\partial x} & \approx \delta_{x-} = \frac{p_{k-1,l,m}^n - p_{k,l,m}^n}{\Delta X} + O(\Delta X) \\
\frac{\partial p}{\partial y} & \approx \delta_{y-} = \frac{p_{k,l-1,m}^n - p_{k,l,m}^n}{\Delta X} + O(\Delta X) \\
\frac{\partial p}{\partial z} & \approx \delta_{z-} = \frac{p_{k,l,m-1}^n - p_{k,l,m}^n}{\Delta X} + O(\Delta X)
\end{align*}
\] (2.26)

The continuous boundary condition Eq. (2.16) in the direction of the negative x-axis may now be approximated using the centered-difference operator Eq. (2.17) in the time-domain, and the first order difference operator in the spatial domain:

\[
\frac{p_{k,l,m}^{n+1} - p_{k,l,m}^{n-1}}{2\Delta T} = -c\xi \frac{p_{k-1,l,m}^n - p_{k,l,m}^n}{\Delta X}
\] (2.27)

A point outside the geometry in the direction of negative x-axis \(p_{i-1,j,m}^n\) may be
written with the discrete boundary condition as

\[ p_{k-1,l,m}^n = \frac{1}{2\lambda\xi} (p_{k,l,m}^{n+1} - p_{k,l,m}^{n-1}) + p_{k,l,m}^n \quad (2.28) \]

By substituting the ghost point into the 3-D discretized wave equation, the following update equation for an x-boundary as illustrated in Figure 2.3 in 3-D is achieved

\[ p_{k,l,m}^{n+1} = \frac{1}{1 + \lambda\beta} [(2 - 5\lambda^2)p_{k,l,m}^n + (\lambda\beta - 1)p_{k,l,m}^{n-1} \]

\[ + \lambda^2(p_{k+1,l,m}^n + p_{k,l-1,m}^n + p_{k,l+1,m}^n + p_{k,l,m+1}^n + p_{k,l,m-1}^n)] \quad (2.29) \]

Where \( \beta = \left(\frac{1}{2\xi}\right) \) and \( \xi \) is the specific acoustic impedance of the boundary. In this case, a simplification is made and for each node a single boundary admittance is used, instead of the component-wise admittances used previously. This may be considered acceptable if the size of the node is small.

---

**Figure 2.3.** Boundary interpretation using the boundary condition with the first-order difference boundary model.

From Eq. (2.29), it can be noticed that the stencil part \((p_{k-1,l,m}^n + p_{k,l-1,m}^n + p_{k,l+1,m}^n + p_{k,l,m-1}^n + p_{k,l,m+1}^n)\) of the equation naturally does not include the node that was eliminated via substitution. To ignore the ghost point from the stencil without actually removing the term from the update equation for air Eq. (2.14), the term has to be forced to be zero during the update. This forcing may be implemented by setting all the pressure nodes lying outside of the geometry to zero before each update.
A free parameter $K$ is introduced to the equation that is used to define how many ghost points are present in each node. With the introduction of $K$, an update that is the same for all nodes of the domain reads:

$$p_{n+1}^{k,l,m} = \frac{1}{1 + \lambda \beta} \left[ (2 - K \lambda^2) p_{n}^{k,l,m} + (\lambda \beta - 1) p_{n-1}^{k,l,m} + \lambda^2 (p_{n+1}^{k,l,m} + p_{n-1}^{k,l,m} + p_{n}^{k,l-1,m} + p_{n}^{k,l+1,m} + p_{n}^{k,l,m-1} + p_{n}^{k,l,m+1}) \right],$$

(2.30)

where $\beta = (\frac{6-K}{2\xi})$ and $\xi$ is the specific acoustic impedance of the boundary node. This scheme is used in simulation of responses in Publication III, IV and V. In Publication I and II, only free-field propagation is considered, and therefore the updates of SRL, CCP and IWB schemes are studied in the absence of boundary conditions.

2.4 Numerical dispersion

The lossless wave equation Eq. (2.7), which is used as the propagation model here admits a plane wave solution of the form:

$$p(\omega, x, y, z, t) = A_0(\omega)e^{-i(\omega t + k_x x + k_y y + k_z z)},$$

(2.31)

where $\omega$ is a real angular frequency and $k_x, k_y$ and $k_z$ wavenumber components, and $A_0$ is the amplitude, $t$ is time, and it is considered that the wavenumber is related to the component values with $k^2 = k_x^2 + k_y^2 + k_z^2$. By inserting the plane wave solution into Eq. (2.7), one arrives to the relation of angular velocity and the wave number

$$\omega = \omega(k) = \pm c k,$$

(2.32)

that is called a dispersion relation. Dispersion relation determines the phase velocity of the propagation, given as

$$c(k) = \frac{\omega(k)}{k}.$$  

(2.33)

Phase velocity defines the speed of propagation of a single wavenumber component. Eq. (2.7) used here is non-dispersive, meaning that the dispersion relation Eq. (2.33) is linear. For the explicit FDTD schemes that are studied in this work, this is not the case. The properties of the phase velocity of the FDTD scheme can be studied by deriving the dispersion relation for the discretized wave equation.

By following the notation of Kowalczyk [68] a dispersion relation for the SRL update Eq. (2.14) for a propagating plane wave may be written:

$$\sin^2 \left( \frac{\omega \Delta T}{2} \right) = \lambda^2 \left[ \sin^2 \left( \frac{k_x \Delta X}{2} \right) + \sin^2 \left( \frac{k_y \Delta X}{2} \right) + \sin^2 \left( \frac{k_z \Delta X}{2} \right) \right]$$

(2.34)
where \( \omega \) is the angular frequency, and \( \hat{k}_x = \hat{k} \cos \theta \cos \phi \), \( \hat{k}_y = \hat{k} \sin \theta \cos \phi \), and \( \hat{k}_z = \hat{k} \sin \phi \) are the numerical wavenumber components where \( \phi \) and \( \theta \) represent the elevation and azimuth angles of the direction of propagation respectively. Eq. (2.34) may be solved for angular frequency \( \omega \) with respect to the wavenumber components:

\[
\omega(\hat{k}_x, \hat{k}_y, \hat{k}_z) = \frac{2}{\Delta T} \arcsin \left( \lambda \sqrt{\sin^2 \left( \frac{\hat{k}_x \Delta X}{2} \right) + \sin^2 \left( \frac{\hat{k}_y \Delta X}{2} \right) + \sin^2 \left( \frac{\hat{k}_z \Delta X}{2} \right)} \right).
\]  

The relative phase velocity can be expressed similarly to Eq. (2.33) as the ratio of the angular frequency and numerical wavenumber:

\[
\hat{c}(\hat{k}_x, \hat{k}_y, \hat{k}_z) = \frac{\omega(\hat{k}_x, \hat{k}_y, \hat{k}_z)}{c \sqrt{\hat{k}_x^2 + \hat{k}_y^2 + \hat{k}_z^2}}.
\]  

The relative phase velocity depends on the values of each numerical wave number component. For the SRL scheme, two cases are of specific interest with regard to the error examination: a plane wave propagating in the direction of a Cartesian coordinate axis, and a plane wave propagating in the direction in which the magnitude of all components are equal, namely the diagonal direction. The case of axis aligned propagation in the direction of positive x-axis, we may write \( \hat{k}_y = \hat{k}_z = 0 \), and for propagation in the diagonal direction an equal wave number component relation reads \( \hat{k}_x^2 = \hat{k}_y^2 = \hat{k}_z^2 = 1/3 \hat{k}_d^2 \). By inserting these values of wave number components into Eq. (2.32), two different expressions for the dispersion relation are achieved:

\[
\begin{align*}
\hat{\omega}_a(\hat{k}_x) & = \frac{2}{\Delta T} \arcsin \left( \lambda \sin \left( \frac{\hat{k}_x \Delta X}{2} \right) \right) \quad \text{(2.37)} \\
\hat{\omega}_d(\hat{k}_d) & = \frac{2}{\Delta T} \arcsin \left( \lambda \sqrt{3} \sin \left( \frac{\hat{k}_d \Delta X}{6} \right) \right) .
\end{align*}
\]

It is evident that the dispersion relation for the axial propagation direction is non-linear. In the case that the simulation is executed at the stability limit of the SRL scheme (\( \lambda = 1/\sqrt{3} \)), the dispersion relation for the diagonal direction reduces to

\[
\hat{\omega}_d(\hat{k}_d) = \frac{\Delta X}{3\Delta T} \hat{k}_d, \quad \text{(2.38)}
\]

which indicates that the dispersion relation is in fact linear in the diagonal direction. Similar analysis for CCP and IWB schemes and for a whole family of FDTD schemes is carried out in [66]. This illustrates an important shortcoming of the SRL scheme and the rest of the studied compact explicit FDTD schemes: the propagation speed of the sound wave in the simulation domain is frequency- and direction-dependent. The direction dependence of the phase velocity in the three schemes is illustrated in Figure 2.4. What can be observed from it is that
Finite-difference time-domain method

each scheme has specific directions in which the phase velocity error is the most severe, and a direction in which relative phase velocity is unity across the wave number component values. The frequency dependence of the wave propagation is such that the high frequencies propagate with a slower phase velocity up to a cut-off frequency, after which the phase velocity becomes complex and the wave components are exponentially decayed [113]. This cut-off frequency directly limits the usable bandwidth of each scheme.

![Figure 2.4](image)

**Figure 2.4.** Relative phase velocity as a function of scaled numerical wave number components. Each row of the plot illustrates a different scheme in the order a) SRL, b) CCP, and c) IWB. First column is the phase velocity in the x-y plane. The second column is the phase velocity along the diagonal of the cubic node volume. The contour lines are at 5% increments in the range of [0.6, 0.95].

Phase velocity is convenient to use as the base of the measurement of dispersion, as it is commonly thought as the speed of sound in room acoustic context. What is often reported as the error of a scheme is the phase velocity error, which signifies the difference between the ideal relative phase velocity of unity, and the relative phase velocity at a given frequency.

As discussed in Publication I, the phase velocity error introduces a frequency,
and distance-dependent \textit{group delay error} to the signal.\footnote{The definition of group delay is the derivative of the \textit{phase shift}, whereas phase delay is the phase shift scaled with the reciprocal of the angular frequency. The difference between the two is that phase delay indicates the delay that is introduced to a single sinusoidal component of the signal, whereas \textit{group delay} indicates the delay introduced to the envelope of a wave packet at a certain frequency [27]. As such, group delay is the more descriptive value here, as it indicates the time-domain spreading of a signal that contains multiple frequency components, such as a transient signal.} Several perceptual studies for the audibility of group delay error have been conducted in different contexts [33, 38, 85]. The effect of the phase velocity error in FDTD simulation has been studied by a few authors [30, 119]. These studies are discussed more elaborately in Chapter 3.4.

\section*{2.5 Computational considerations}

The computational aspects of FDTD simulation have had a fair amount of interest in the research community. A major driver for the interest has been the increased availability and performance of parallel computing hardware such as \textit{graphics processing units} (GPUs), the usage of which outside of graphics-related computing is fittingly termed \textit{general-purpose computing on graphics processing units} (GPGPU). Many of the proposed compact explicit FDTD schemes are possible to implement as data parallel algorithms, and therefore usage of GPGPU to accelerate the execution is straightforward. The performance benefit of the usage of GPGPU is widely agreed on in the community, and therefore some details of using GPU hardware are discussed here.

GPUs are devices originally dedicated to graphics processing for gaming, rendering and visualization applications, where a large number of geometric primitives are rendered to a two or three dimensional medium within a given frame rate limitation. [35] The usage of GPU devices for other uses became increasingly popular after the first GPGPU \textit{application programming interfaces} (APIs) were published [87, 122]. GPUs have been found to be suitable for several different audio applications, such as large-scale convolution [140, 108, 9], and sound synthesis [146, 108]. The main requirement for an efficient use of a GPGPU is that the computation can be executed in parallel for a large number of data points that are not dependent on the values of each other.

First study proposing the usage of graphics processing units for wave-based acoustics simulation was conducted by Röber et al. [99]. In the study, the simulation update was implemented using texture objects and fragment shaders without the usage of designated GPGPU APIs. The first comprehensive study on the implementation of FDTD schemes on GPU hardware using the CUDA programming language was published by Savioja [102]. Savioja conducted several experiments with two different explicit FDTD schemes and various boundary conditions. [107] Further studies have included the introduction of an uniform update using first-order accurate boundaries [138], different
Finite-difference time-domain method

memory access mechanisms [79], investigation of higher-order schemes with larger stencils [132] [51], and the possibility of using implicit schemes with GPU hardware [50].

The main performance limitation to FDTD simulation with GPUs is the data bandwidth between the actual processing units on the GPU and the memory units on the GPU [102, 51]. The limitation can be described with a parameter called compute to global memory access (CGMA) ratio [64, p. 96], which indicates the ratio of the performed floating-point operations and global memory accesses. An example similar to what is presented by Kirk et al. illustrates the implications of the ratio: for a GPU device with a memory bandwidth of 288 GB/s, and capability of 1,680 giga floating-point operations per second (GFLOPS), and a compute kernel CGMA of 1.0, the kernel may reach a performance of only \( \frac{288}{4} = 72 \) GFLOPS. This value is only a fraction of the peak 1,680 GFLOPS, and therefore better performance may be achieved by improving the kernels CGMA. For compact explicit FDTD schemes, the CGMA ratio is close to 1.0. Modern GPU architectures have included dedicated cache hierarchies in overcoming bandwidth limitations in some cases [62]. However, due to the large number of memory reads and writes involved in the FDTD simulation update, ideal cache utilization is not attainable [51]. A good review on computational techniques for FDTD simulation in general may be found in the work of Webb [139].

![Diagram of partition scheme](image)

**Figure 2.5.** Illustration of the partition scheme used in the communication between the different devices of the distributed system. The color indicates the common layers between neighboring partitions called halo layers. After a simulation update is executed in parallel for each partition, the halo layers are exchanged between neighboring partitions.

A distributed approach is a viable solution to scaling up the simulation do-
Finite-difference time-domain method

main size for off-line simulation. The distribution and synchronization of the simulation domain for a compact explicit FDTD simulation is fairly simple to implement. As only immediate neighbors are involved in the simulation update step of a compact scheme, a single shared layer is needed between each device-device interface. Such a layer is commonly referred to as a *halo layer*. The partition and the halo layers are illustrated in Figure 2.5. The synchronization of the domains is done by exchanging halo layers between adjacent domain partitions after each simulation update.

For real-time use, some limitations emerge due to synchronization latencies and the limited memory bandwidth between the devices involved. The use-case of distributed real-time simulation for a VR and *augmented reality* (AR) application was studied in Publication III.
3. Experimental methods for perceptual evaluation of room acoustic simulation

The perceptual evaluation of audio in general can be considered to contain two levels [7, p. 3]: perceptual measurement and affective measurement. The difference between the two is that perceptual measurement aims to measure the sensory strength of a chosen auditory attribute, whereas affective measurement aims to measure the overall impression of an auditory stimulus, possibly with no assumptions on the auditory attributes. Affective measurement therefore includes all variables regarding the participant’s mood, personal taste, recent experiences and other such subjective factors. Such measurement may be considered excessively variable and possibly difficult to replicate, and are conducted for large panels in order to acquire information about patterns at the population level [144, pp. 62 – 63].

On the contrary, perceptual measurement leans on the assumption that some perceptual characteristics of a sensory experience are rated in a similar manner across the individuals of a panel of assessors. For such measurement, a crucial step is to define, or possibly elicit the attributes that are used to quantify the sensory experience. For the evaluation of audio, several sets of attributes and recommendations for developing attributes have been suggested [71, 57]. It is assumed that the sensory strengths of each attribute in stimuli or the relation of the sensory strengths between stimuli can be measured using psychometric techniques.

In terms of perceptual quality of acoustic simulation, some questions that have commonly appeared in studies include: “Is the simulated sound field perceptually similar to a reference sound field?” [77, 78, 145, 39, 6, 93], “Does the simulated sound field contain audible artifacts in comparison to a reference?” [30, 90, 119], “How well does the simulated sound field correspond to the presented visual stimuli?” [110, 3], and “How plausible is the presented condition is in a holistic sense?” [3]. Each of these questions can be studied in terms of multiple attributes with the exception of the last, which can be considered an affective measurement. For a single attribute, the interest of the experimenter is often the question of whether a participant can discriminate between two stimulus levels, or examining how different stimuli levels correspond to one another. These questions may be addressed using discrimination testing [74, pp. 79 –99], and scaling
methods [7, pp. 69 – 92] respectively. In the case that the stimuli levels can be controlled explicitly, which is often the case in simulated responses, a series of discrimination tasks may be used to acquire sensory thresholds for different attributes and in doing so, find generalizable evidence on how the method of choice behaves in a perceptual sense.

In this chapter, some general techniques related to measurement of perceptual qualities of room acoustic simulation are reviewed. First, an overview regarding how acoustic conditions are commonly presented to a test subject is given. Then, the sensory methods used in publications I, II, and V are given more detailed overview. As the scope of the thesis is limited to perception of numerical error, the discussion is limited to perceptual measurement, although similar techniques, such as the approaches of direct and indirect scaling that are discussed in Section 3.3, may be used for affective measurement. Finally, some methods utilized in previous studies on the perceptual quality of room acoustic simulation techniques are discussed in brief.

### 3.1 Auralization methods

A major decision in any listening experiment is the determination of how to present the stimuli to the participant. The question becomes even more central when the presented stimulus represents a room acoustic condition, as different auralization techniques may yield different results. [120, 125]

The term *auralization* is defined by Kleiner et al. [65] as follows:

> Auralization is 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.

A more compact and broader definition for the term is proposed by Vorländer [134, p. 103]:

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

The definition given by Vorländer is broad enough to be confused with *sonification*, which is commonly used for creating an audible representation of any dataset, not only data representing an acoustic system or a sound field. The definition given by Kleiner et al. is closer to the etymology of the word auralization (Latin *auris*, ear), as the process is related to how humans perceive sound, and also contains the notion that the sound source is located in a space. The definition of Kleiner et al. limits the term to modeled responses, which is not the case with the definition of Vorländer, in which measured and synthesized datasets are also included. In this work the definition by Kleiner et al. is used.
with the addition that measured data representing an acoustic system may be used as the basis of an auralization.

Perhaps two of the most common approaches to auralization in room acoustic research are reproduction using binaural technique via headphones [84], and playing back multichannel audio via a loudspeaker array. The setup of the loudspeaker array may be configured according to different standards, but here loudspeaker array is assumed to be able to play back sound both in horizontal and vertical directions at the listening position, namely the system is assumed to be periphonic [44]. In a physical sense, the aim of both methods is to reproduce the sound field at the subject’s eardrums, as it would emerge in a corresponding scene in a real environment. Often such a strict physical criterion is considered unattainable or unnecessary, and auralization methods are compared according to the perceived quality. In a perceptual sense, the sound field would not need to be identical but it should result in the same sensory experience. What is considered to lead to similar sensory experience [83], is being able to reproduce certain, if not all, localization, as well as environmental cues [100, pp. 10 – 33].

A common assumption in room acoustics is that the acoustic system including the source, the enclosure, the medium and the receiver, for discrete source and receiver locations, can be described with an impulse response. A single impulse response is typically an array of pressure values as function time, and as such does not contain any directional information. As the localization and environmental cues largely depend on directional information, it is common in practice that a set of impulse responses at the proximity of the receiver position is measured. By using the location information of the different receiver positions and temporal information in the corresponding impulse responses, directional characteristics of the sound field at each time point in the proximity of the receiver position can be estimated. This process is referred to here as spatial encoding.

In the following, the two typical auralization approaches are briefly reviewed and an overview of the associated spatial encoding techniques is given.

### 3.1.1 Binaural technique

The binaural technique aims to directly reproduce the two signals at each eardrum of the subject, as they would emerge in a corresponding real sound field. [84] What this approach entails is that all the head-related effects introduced to the sound field need to be addressed with either signal processing techniques, or by directly recording the pressure inside the subject’s ears. The latter approach is commonly referred to as binaural recording, and the former binaural synthesis. An alternative for a binaural recording with a subject individual is a dummy head recording, where the recording is made using microphone capsules that are mounted in the ears of a model of human head. Room impulse responses that are measured with either probe microphones inside the subjects ears, or with a dummy head, are referred to as binaural room impulse response (BRIR).
BRIR could be considered a form of spatial encoding scheme, as the directional information is included in the two signals in the form of the direction dependent level, time and spectral differences.

In the case that the signal processing approach is chosen, some precursory preparations need to be carried out in order to apply the binaural technique. To introduce the effects related to the sound field’s interaction with the subject’s head, a set of the head-related transfer functions (HRTFs) of the subject are measured. HRTFs are a set of transfer functions that represent the propagation of sound from individual directions to each ear of the subject in free-field. HRTFs can be measured using similar microphone techniques as in binaural recording, and specialized machinery that enables rapid source position changes to cover all the directions intended for measurement. Due to the laborious nature of the measurement procedure and the specialized hardware involved, public databases [43, 1] for both human and dummy head HRTFs are often used in research. An alternative approach to measurement is to approximate the subject’s HRTFs with simulation methods [46, 55]. In such case, a 3-D model of the subject’s head need to be constructed using scanning techniques. Naturally, the quality of the mesh has been identified as an important factor in determining the quality of the simulated HRTFs [147].

To include the room acoustic condition in the binaural synthesis, a spatially encoded response of the room condition is processed with the HRTF data. Depending on the encoding scheme, the room acoustic data may be combined with the HRTF data using function decomposition such as spherical harmonic (SH) representation, or if direction of arrival (DOA) information for different temporal field components is a part of the encoding scheme, HRTFs corresponding to each direction can be used as such.

To accommodate the fact that the head of the subject may be moving with respect to the intended room acoustic condition, head tracking is commonly added to the binaural synthesis system. The importance of head tracking, environmental cues, and individualization of HRTFs to binaural synthesis has been an interest of research for several decades [8].

In Publication I and II, headphone reproduction was used, but the binaural technique was not employed as such. The stimuli presentation was done using diotic playback over headphones that contained a diffuse field compensated response. Diotic playback means that the same signal is played back to both ears. Such a case could emerge when a sound source is located on the median plane of the listeners. Binaural technique was not implemented as the playback method in Publications V, as the author considered the different aspects of the technique, such as individualization of the HRTF and individualization of the headphone compensation filter important yet prone to error in the acquisition. Loudspeaker reproduction was used to have the conditions represented in a uniform manner to the participants.
3.1.2 Loudspeaker reproduction

Loudspeaker reproduction in a periphonic sense consists of a fixed number of loudspeakers that are set up to physically surround the subject both in the horizontal plane and at different elevations. Commonly, the listening room is acoustically treated to minimize the interaction with the sound field that is produced by the loudspeakers. One of the main advantages of loudspeaker reproduction is that all head-related effects are inherently included as the reproduced sound field surrounds the head of the subject. There are several approaches to construct the signals that are played back with the loudspeaker array which depend on the type of spatial encoding that is used to represent the measured or simulated acoustic field.

In the case that the encoding scheme contains DOA estimates for some assumed directional field components, such as plane waves in certain methods, each directional field component is commonly panned between a subset of loudspeakers using local amplitude panning laws, such as vector based amplitude panning [94], or with global panning laws, such as ambisonic panning [80], where the directional field component is encoded using SH representation, and then decoded to the full loudspeaker array. Combinations of the two approaches have been proposed [148]. Very simple approaches, such as nearest-loudspeaker reproduction, where each directional field component is played back from a single loudspeaker that is located closest to the DOA of the component [88] have been shown to preserve some aspects of reproduced concert hall acoustics better than amplitude panning. If the spatial encoding scheme makes estimates for the diffusiveness [83] of a certain temporal part of the sound field, the diffuse portion is played back from all of the loudspeakers, or from a subset of the loudspeakers using decorrelation techniques.

There are a few problematic properties in direct use of panning techniques. One is that the temporal distribution and the relative loudness of the reproduced directional field component change if the listener moves with respect to the loudspeakers. Therefore the listening area where the reproduction is considered accurate is small and often referred to as the “sweet spot”. Another is that amplitude panning may lead to variable spread of the location of the positioned directional field component depending on the number of loudspeakers that is used in the array, and what is the aperture angle between the individual loudspeakers used in the panning [92, pp. 50 – 51].

A more physically motivated approach that is capable of overcoming the problem of small listening area is wave field synthesis (WFS), which aims to reconstruct a given wave field inside of a given volume.[11] The approach is based on the Huygens’ principle, which states that any point of a wave front can be considered an individual monopole source. In practice, WFS is implemented using a large number of densely distributed loudspeakers with which different types of waveforms can be produced over the reproduction area using a set of driving functions. A practical limitation of the method is that at frequencies for
which the inter-loudspeaker distance is half of the wavelength or less, spatial aliasing will occur and introduce inaccuracies to the wave field. A comprehensive review of the principles of loudspeaker reproduction techniques can be found in the work of Spors et al. [120].

3.1.3 Spatial encoding

Both binaural synthesis and loudspeaker array reproduction aiming at reproducing room acoustic conditions need information about the direction from which a wave front, or a group of wave fronts, should be reproduced to the subject at each point in time. Several methods to encode this spatial information have been proposed.

A method to estimate and encode the directional information from a set of impulse responses into a set of image source locations and amplitude values, the *spatial decomposition method* (SDM), was proposed by Tervo [125]. The method processes impulse responses measured at arbitrary receiver positions in consecutive, overlapping time windows, and estimates a DOA for a single plane wave for each window using cross-correlation and the inter-receiver distance information. The result of the encoding scheme is a single pressure value and a DOA estimate per sample. A more advanced method to estimate the directions of arrival for multiple plane waves per time-frequency analysis window using spherical microphone array was later introduced [124].

The ambisonic framework [44] encodes the directional information in first order SH coefficients. A practical convenience of the encoding scheme is that the encoded signals can be directly measured using a coincident microphone setup [45]. When impulse responses to a sufficiently large number of receiver positions are available, a spatial encoding to a higher-order of SH coefficients in a form of plane wave decomposition [97] is possible.

Spatial encoding schemes that are designed to take into account the temporal and spectral resolution of human hearing include *spatial impulse response rendering* (SIRR) [83], and *directional audio coding* (DirAC) [95]. Both methods aim to decompose the signal into directional and diffuse components using time-frequency analysis. SIRR, as the name suggests, uses this approach in auralization of room impulse responses, whereas DirAC may be considered more general technique.

Regarding the spatial encoding of different simulation methods, GA methods possess a convenient property that the DOA of each reflection path is a part of the numerical solution itself. This property makes the final auralization absent of the error that may be introduced with the spatial encoding, and GA solutions have therefore been used as a reference for the perceptual evaluation of spatial encoding schemes [96, 125].

Spatial encoding techniques for measured responses and wave-based simulation responses can be considered conceptually similar, although several practical differences can be identified. In the case of room acoustic measurements, the
number of microphone capsules that are used is often limited (commonly fewer than hundred). Additionally, the position of each capsule is known with limited accuracy. For simulation, the number of receiver locations where the responses are solved is limited by the memory on the device that is used for the calculation. The receiver positions in the simulation are explicitly set and can be considered accurate and therefore reduce inaccuracies in spatial encoding. Due to these properties, a spatial encoding scheme proposed by Scheaffer et al. [116] using higher-order SH coefficients was used in Publication V. An overview of this scheme is presented in the following.

Plane wave decomposition using spherical harmonics

A common approach to spatial encoding is to use SH representation. For audio use, SH representation is often used synonymously with ambisonic [44, 37] representation, which is essentially a SH representation of audio signals, with the addition of technical considerations for recording techniques.

The method proposed by Scheaffer et al. [116] uses a set of impulse responses measured at multiple positions within a region of an open sphere centered at the intended receiver position to generate a plane wave decomposition (PWD) [97] of a sound field. The PWD is acquired using SH representation that is based on spherical Fourier transform, synonymous with spherical harmonic transform (SHT), where functions that are defined on a unit sphere are presented as a weighted sum of a set of basis functions. The basis functions are termed spherical harmonics, and are defined as

\[ Y_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos \theta)e^{im\phi}, \] (3.1)

where \( m \) is the degree, and \( n \) is the order of the function. The SHT and the inverse transform ISHT is defined as

\[ f_{nm} = \int_0^{2\pi} \int_0^\pi f(\theta, \psi)Y_n^m(\theta, \phi)\sin \theta \, d\theta \, d\phi = \text{SHT}\{f(\theta, \phi)\}, \] (3.2)
\[ f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} f_{nm}Y_n^m(\theta, \phi) = \text{ISHT}\{f_{nm}\}. \]

where \( f(\theta, \phi) \) is a continuous function on a sphere and \( f_{nm} \) are the SH coefficients for the function. If the assumption is made that the sound field can be decomposed into a set of plane waves, it may be represented with a plane wave density function [97] \( a(k, \theta, \phi) \), where \( k = 2\pi f/c \), \( k \) is the wavenumber, and \( f \) is the frequency. The SHT of the pressure on a sphere using the plane wave density function can be calculated using Eq. 3.1:

\[ p_{nm}(k,r) = \int_0^{2\pi} \int_0^\pi a(k,\theta,\phi)b_n(kr)Y_n^m(\theta,\phi)\sin \theta \, d\theta \, d\phi \]
\[ = b_n(kr)a_{nm}(k) \] (3.3)
Experimental methods for perceptual evaluation of room acoustic simulation

where $b_n(kr)$ is a radial function that relates the pressure values on a sphere to the SH representation.

In the case of a FDTD simulation, the pressure is solved at each discrete location within a domain. Considering $Q$ grid nodes that are captured within a chosen spatial region with a center point $c_q$, having a radial distance of $r_q$ from the center point and is located in the direction $\theta_q, \phi_q = \Omega_q$ from the point $c_q$. The captured pressure at each point can be expressed in a vector form in the frequency domain as

$$p(r, \Omega, k) = [p_0(r_0, \Omega_0, k), p_1(r_1, \Omega_1, k), \ldots p_Q(r_Q, \Omega_Q, k)]^T,$$  
(3.4)

from which an expression for pressure as a function of radial distance and direction may be written as an approximation using Eq. 3.2

$$p(k, r_q, \Omega_q) \approx \sum_{n=0}^{N} \sum_{m=-n}^{n} a_{nm}(k)b_n(kr_q)Y_m^n(\Omega_q).$$  
(3.5)

A matrix notation for the SH coefficients that are weighted with radial functions can be written as:

$$B^T = \begin{bmatrix}
  b_0(kr_1)Y_0^0(\Omega_1) & \ldots & b_0(kr_1)Y_0^0(\Omega_Q) \\
  b_1(kr_1)Y_1^{-1}(\Omega_1) & \ldots & b_1(kr_1)Y_1^{-1}(\Omega_Q) \\
  \vdots & \ddots & \vdots \\
  b_N(kr_1)Y_N^N(\Omega_1) & \ldots & b_N(kr_1)Y_N^N(\Omega_Q)
\end{bmatrix},$$  
(3.6)

and the SH coefficients for the plane wave density function can be expressed in a vector form as:

$$a_{nm} = [a_{00}, a_{1(-1)}, a_{10} \ldots a_{NN}]^T.$$  
(3.7)

With this notation, Eq. (3.5) for pressure can be written as:

$$p = Ba_{nm},$$  
(3.8)

from which the plane wave density function coefficients may be approximated as:

$$a_{mn} \approx \hat{a}_{nm} = B^+p,$$  
(3.9)

where $B^+$ is the Moore-Penrose pseudoinverse of $B$. Additionally the regularization of the radial functions that is proposed by the authors of the approach [116] is employed in Publication V. The regularization is done to ensure that the matrix $B$ is well-conditioned. In practice this means that the values of the radial functions $b_n(kr)$ are limited to have a maximum allowed attenuation.

In Publication V the loudspeaker signals are constructed from the plane wave density function by extracting time-domain impulse responses at azimuth and elevation angles at the range $[-180, 180]$ and $[90, -90]$ degrees, respectively at
Experimental methods for perceptual evaluation of room acoustic simulation

increments of 3 degree. These impulse responses, representing plane waves propagating from each respective direction towards the receiver position, are panned using the nearest-loudspeaker method [88] for the considered loudspeaker array.

3.2 Threshold measurement

Early models for how the human sensory system detects physical stimuli included Weber's law, and Fechner's law [36], which both suggest that the ability of humans to tell the difference in magnitude between two stimuli depends on the level of the stimuli itself. Weber formulated the hypothesis into a simple theory that the ratio of the perceivable intensity difference and the intensity is a constant, whereas Fechner further developed Weber's law into a logarithmic relationship between the sensation intensity and the physical stimulus intensity. The study of the relation between physical stimuli and perceived sensory experience was termed psychophysics by Fechner.

The origin of threshold measurement in psychophysics stems from the work of Fechner with the establishment of method of limits and method of constant stimuli [20, pp. 36 – 37]. The absolute threshold was considered the stimuli level that produced 50% probability for the participant to report that stimulus was present. This point was measured either with ascending and descending stimulus levels in the method of limits, or by measuring the frequency of responses for different stimuli levels in the method of constant stimuli. With the method of limits, the two threshold values for ascending and descending stimuli levels were typically different, and a mean value of the two was considered the threshold. Another threshold value that was proposed is the difference threshold, which is determined by increasing or decreasing the stimulus level from and to a chosen reference level, and recording the level when the participant responds that the two stimulus levels are perceived as either differing or equal respectively. Difference threshold is also referred to as just noticeable difference (JND) in practice.

Since the early techniques, a large number of methods have been introduced to model and measure sensory experience as a function of physical stimuli. In the following, several concepts are reviewed with regard to threshold measurement in practice.

3.2.1 Psychometric function

The frequency of responses, whether or not a certain sensory attribute is present, varies for different stimulus levels. The function that indicates the probability of response as a function of stimulus level is called the psychometric function. Initially, a cumulative distribution function of the normal distribution function was used to represent the psychometric function [20, p. 40], but several different types of functions have been used since [130]. A common alternative is a
Experimental methods for perceptual evaluation of room acoustic simulation

psychometric function based on Weibull distribution [141], which is used in the following form in the software package [89] and subsequently in the experiments in Publications I, II and V:

$$\Psi(x, T)_{\text{weibull}} = \zeta \gamma + (1 - \zeta)(1 - (1 - \gamma)\exp[-10^\beta(x-T)]) \quad (3.10)$$

where $\zeta$ is the lapse rate, $\gamma$ is the guess rate, $\beta$ is a shape parameter indicating the steepness of the function and $T$ is a threshold parameter controlling the individual probabilities associated with each stimulus intensity level.

Two notable properties about the shape of psychometric function should be made. The task given to the test subject determines the probability of giving a positive response by guessing. For example, for three-alternative forced-choice trial, where the participant is tasked with distinguishing between stimuli from amongst three alternatives including two reference samples, the probability of guessing correctly is $p = 1/3$. Similarly, in two-alternative-force-choice trial the probability of guessing the correct response is $p = 1/5$. This so called guess rate, or rate of false positives, which determines the lower horizontal asymptote of the psychometric function. Similarly, a rate for false negative is usually determined for cases where the participant, via motoric error or due to memory lapse, gives a false response in a case where the stimuli is considered very easy to discriminate. This rate determines the higher horizontal asymptote of the psychometric function and is referred to as the lapse rate. Slightly different variations of the function Eq. (3.10) are presented in [137, 63], which introduce the offsets regarding guess and lapse rates with a different arithmetic.

### 3.2.2 Adaptive procedures

The method of constant stimuli may be considered laborious when the interest of the experimenter is to measure a specific probability level of positive response, which corresponds to a single point on the psychometric function. For such a task, an adaptive procedure - a procedure where the experimental condition for each trial is determined by the result of the previous trial - can be significantly more efficient. The general aim for adaptive methods is to measure the subject’s sensory abilities in the shortest possible time without sacrificing accuracy.

The first statistical analyses of simple adaptive procedures, such as the up and down method, were introduced in the middle of 20th century [4, 34] for measuring the sensitivity of explosives. In psychoacoustics [133] and psychophysics in general [31], the up and down method consists of sequential trials, where the task for the participant is to indicate whether a stimulus is audible. After a positive response, a lower stimulus level is presented in the following trial, and a higher stimulus level in the case of a negative response. This procedure measured the 50% probability of perceiving the stimulus with a yes-no task. Different probabilities of a correct answer can be measured using transformed up-down procedures [76] where the change in the stimulus level is conducted after a number of positive or negative responses are recorded. For example a
transformed up and down rule “2-down 1-up”, is where the stimulus level is reduced after two consecutive positive responses and increased after a single negative response.

Up and down, or staircase methods can be described as non-parametric [130], as no parametric model for the shape of the psychometric function or the distribution of the threshold value are made, except that the psychometric function is monotonic. The efficiency of an adaptive procedure may be improved by using a parametric model for the psychometric function, such as Eq. (3.10). In these procedures, the information acquired from the trials is used to fit a psychometric function in respect to a threshold parameter at a certain probability, the slope parameter associated with the used psychometric function or both. The parameterization can also be used to decide the next stimulus intensity level, which is discussed in the following section. Several different maximum likelihood and Bayesian approaches have been proposed and interested reader may find thorough reviews of adaptive methods by Treutwein [130] and Leek [75].

**QUEST procedure**

In Publications I, II and V, the QUEST procedure was used to measure threshold values for dispersion error in various conditions. QUEST was proposed by Watson and Pelli [137] as a method to measure sensory thresholds using Bayesian approach in taking all previous trials into account while calculating the next stimulus intensity level. The reasoning behind using this method is that it is shown to be efficient [63], and that a ready package based on the implementation by one of the authors of the method was available [26] [89].

The method uses assumptions for both the shape of the distribution of the individuals threshold, and the shape of the psychometric function, which are:

- The psychometric function has the same shape under all conditions when expressed as a function of log intensity. The position on the intensity axis is determined by the parameter $T$, which is the true threshold of the participant

- The participants’ true threshold $T$ does not vary from trial to trial

- Individual trials are statistically independent

The method assumes an a priori distribution for the threshold value at the chosen probability level, and a parametric representation of the psychometric function. Two sources of information are used to estimate the threshold: the a priori distribution, and the data that is acquired from the trials. The a priori distribution and the shape of the psychometric function may be chosen by the experimenter. The **probability density function** (PDF) of the a priori distribution is referred to here as $t(\theta_{\text{prior}})$ and is a function of the initial threshold estimate $\theta_{\text{prior}}$.

After $n$ trials, the intensity levels that were presented can be listed as a
Experimental methods for perceptual evaluation of room acoustic simulation

vector \( \mathbf{x} = (x_1, \ldots, x_n) \). The estimate of the true threshold is indicated with \( \theta \).

As the assumption was made that the psychometric function follows a chosen parameterization, the probability of positive and negative response \( p_+ \) and \( p_- \) for a given intensity level can be expressed with the chosen psychometric function \( \Psi(x_i, \theta) \) and its complementary:

\[
f(x_i|\theta) = \begin{cases} p_+(x_i, \theta) = \Psi(x_i, \theta) \\ p_-(x_i, \theta) = 1 - \Psi(x_i, \theta) \end{cases}
\] (3.11)

Now by combining the information, the posteriori PDF for the threshold parameter after the first trial may be written as:

\[
p(\theta|x) = \frac{f(x_1|\theta)p(\theta)}{h(x_1)}
\] (3.12)

The function \( h(x) \) is the prior PDF of the data, which can be evaluated with
\[
h(x) = \int_{-\infty}^{\infty} f(x|\theta)p(\theta)\ d\theta.
\] It is considered only for normalization, and therefore it can be excluded in most cases, as it does not affect the shape of the posteriori PDF that is a function of \( \theta \).

In other words, the posteriori distribution is achieved by weighting the a priori distribution with a likelihood function, which in this case is the psychometric function and its complement. The QUEST procedure progresses from trial to trial by substituting the prior distribution with the previous posteriori distribution after each trial

\[
p(\theta|x) = f(x_i|\theta)p(\theta|x_{i-1}).
\] (3.13)

Watson and Pelli suggest that the next intensity level is chosen to be the mode of the posteriori PDF, which is the maximum likelihood estimate of the threshold. King et al. [63] proposed that the intensity level should be chosen to be the mean of the posteriori PDF. This approach minimizes the expected variance of the threshold estimate after the following trial. The mean of the posteriori PDF is used as the intensity level of each trial in the procedure used in Publications I, II, and V.

3.3 Scaling

As threshold measurements commonly address the question at what level the participant can discriminate between a stimuli and a reference, the approach inherently does not give information about how different levels correspond to one other. Experiment designs where the participant is given a task to rate or order the stimulus according to a specified dimension can provide such information. Techniques for these experiments are commonly called scaling procedures.

Scaling may be divided into two main categories [7, p. 68]: direct, and indirect scaling. Direct scaling involves a task for the participant where a stimuli is placed into a category (nominal scale), sorted into a specific order along a
dimension (ordinal scale), given a score on an interval (interval scale), or rated in proportion to a reference (ratio scale). Several different interval scales have been proposed for the assessment of audio quality [61, 58, 59], and, for example, the popular MUSHRA [59] method falls into the category of direct interval scaling.

Direct scaling can be problematic when a stimulus is considered to be multidimensional in perceptual sense, namely that there are multiple sensory attributes that may be used to describe a particular stimulus. Rating multidimensional stimuli along a single axis may lead to differing interpretations among the subjects. This can lead to a violation of stochastic transitivity: the participant tasked to rate three stimulus, A, B and C, may rate the stimuli so that A is preferred over B, B over C, and C over A. This phenomenon is not captured using direct scaling. Additionally, the evaluation of the sensation magnitude among the subject panel can differ, and different rating strategies may introduce a bias to the responses.

Indirect scaling involves a number of tasks in which the stimulus is judged in comparison to reference, or another stimulus in magnitude along a given sensory dimension. The assumption in indirect scaling is that the number of times a stimulus is judged to be greater in magnitude on the scale is related to the magnitude of the sensory difference between the stimuli. The downside of the approach is that if the discrimination probability associated with the comparison approaches 100%, the scale value difference becomes undefined [74, p. 167], and therefore indirect scaling is more appropriate for small sensory differences.

3.3.1 Paired comparisons

A popular alternative to evaluating room acoustics conditions in terms of various attributes or preference has been paired comparisons [115, 5, 28, 29, 126], and it is a part of an ITUR recommendation [60] as a method of subjective measurement of sound quality. Paired comparison belongs to the category of indirect scaling methods. The assumption for the paired comparisons procedure is that the scale value of a sample within a group of samples is proportional to the frequency of it being preferred over other samples in terms of a given attribute. The product of a paired comparisons experiment is a matrix containing the times each condition is preferred over another.

The Bradley–Terry–Luce (BTL) model is used in Publication V for estimating the rating values of the different test conditions from the paired comparison data. As described by Bradley [25], the paired comparison method assumes that each presented condition from $n$ conditions has a true rating $[u_i, ..., u_n]$. The ratings have three properties: every rating $u_i \geq 0$, $\sum u_i = 1$, and that from two conditions $i$ and $j$, the probability that condition $i$ is chosen is $\pi_{ij} = \frac{u_i}{u_i + u_j}$. As the comparison is always a binary choice, the probability of achieving $N_{ij}$ number of judgments that the condition $i$ being chosen over $j$ can be written in
terms of binomial distribution:

\[
\binom{N}{N_{ij}} \pi_{ij}^{N_{ij}} (1 - \pi_{ij})^{N - N_{ij}} = \binom{N}{N_{ij}} \pi_{ij}^{N_{ij}} \pi_{ji}^{N_{ji}},
\]

from which a likelihood function for the observation containing all pairs can be derived:

\[
L = \prod_{i < j} \pi_{ij}^{N_{ij}} \pi_{ji}^{N_{ji}},
\]

where \(N_{ij}\) is the number of times option \(i\) is preferred over option \(j\) from \(N\) comparisons, and \(\pi_{ij}\) is the probability associated with the preference. A log likelihood may be derived directly from Eq. (3.15):

\[
\log L = \sum_{i < j} N_{ij} \log(\pi_{ij}) + N_{ji} \log(\pi_{ji}).
\]

The log likelihood function can be written in terms of the ratings:

\[
\log L = \sum_{i < j} N_{ij} \log \left( \frac{u_i}{u_i + u_j} \right) + N_{ji} \log \left( \frac{u_j}{u_i + u_j} \right).
\]

Optimal rating values that maximize the log likelihood function can be found using numerical optimization routines. A MATLAB package implemented by Wickelmeier and Schmid [143] was used in Publication V to solve the rating values from the comparison matrix.

### 3.4 Previous experiments

Now that some general approaches of perceptual measurement and auralization techniques have been reviewed, a brief review on the past work regarding the perceptual evaluation of simulation methods is given. Several authors have approached the sensory evaluation problem with different methods. In terms of the two levels of perceptual evaluation, the evaluations are mostly conducted using perceptual measurement, although some studies include questions typical of affective measurements.

Lokki et al. conducted several experiments [77, 78] concerning the perceptual quality of auralizations generated with the DIVA system [106]. The first study [77] consisted of a similarity-rating task of the simulated and recorded BRIRs using four different attributes on a five-point scale. The used attributes were sound source localization, externalization, sense of space, timbre, and the scale ranged from very different to very similar. The test consisted of six conditions: two static listener positions, two stimulus including a head turn, and two “walk through” responses. The second experiment [78] consisted of a discrimination task combined with an interval-scaling task between simulated and recorded conditions. The participant was asked to identify from two options A and B,
which one is equal to a reference, X, and then rate the difference between the remaining sample and the reference. The difference rating was given on a 40-interval scale describing the difference ranging from very annoying to imperceptible. Four different room conditions were used: two listening positions and inclusion/exclusion of first-order diffraction effects. The general conclusions of the two experiments were that the simulated responses were easier to differentiate from the recorded when a transient-like stimulus was used, that the simulated responses were in some cases considered very natural sounding, and that diffraction modeling was not improving the perceived quality in all cases.

Zahorik [145] conducted an experiment where the similarity of a set of simulated and recorded individualized BRIRs were measured using multidimensional scaling (MDS) techniques. Simulations were carried out using GA methods for early reflections and late reverberation was synthesized. BRIRs were measured for each participant in a single room condition, but several different simulated room conditions were included in the study. Some of the simulated room conditions were designed to closely resemble the room where the measurements were carried out, and several other room conditions were different in size and surface properties. The task for the participant was to rate on a 100-point scale how similar each room condition pair was perceived. It was concluded from the results of the MDS analysis that the similarity rating can be reduced to two perceptual dimensions that correlate with reverberation and sound spaciousness respectively. Interestingly, the simulated condition closest to the group of measured conditions in the analysis result was a BRIR where only early reflections were modeled. The results provided evidence to that timbre did not contribute to the perceptual similarity as an independent factor, and that similarity ratings with individualized and non-individualized BRIRs appeared very close to each other in the results.

Pelzer Vorländer [90] conducted a listening test measuring the perceptual effect of geometry simplification in GA simulation. The method of constant stimuli was used with eight different geometric model conditions. The geometric model conditions had a varying level of detail. The level of detail was controlled by adding structures with a decreasing size into the geometry. The minimum structure sizes associated with each condition were 10, 5, 2, 1.25, 0.6, 0.3, 0.15 and 0.07m. The model with the highest level of detail therefore had structures of each size, whereas the model with the lowest level of detail had only structures greater than 10 m in size. Each condition was compared with the model with the highest level of detail using a three-alternative-forced-choice task. The task presented to the participant was to identify which auralization was different amongst the three alternatives, from which one was the stimulus and two were references. The method of the stimuli presentation was not reported. A Gaussian cumulative distribution function was fitted to the correct discrimination frequency data as a function of the logarithmic detail level. It was concluded that at a 75% probability level of a correct discrimination was at a detail level of 0.7 m.
Foteinou and Murphy [39] measured the similarity of reverberation between simulated and measured impulse responses on a 10-point scale that represented the difference from very similar to very different. The authors conclude that the participants were able to discriminate between the measured and simulated responses and tended to rate the simulated responses as “more reverberant”, contradicting the comparison of objective parameters. The authors speculate that the result may have been affected by the choice of model parameters.

Aretz et al. [6] conducted several preliminary listening tests as a part of a comprehensive objective evaluation study of a combined FEM and GA simulation. Simulated and measured BRIRs of a car passenger compartment and a studio control room were compared. Participants were asked elaborate on the perceptual differences between the measured and simulated responses in their own words. It was concluded that in general, the participants were able to discriminate the simulated responses from the measured responses although no audible artifacts were reported in the simulated responses. The use of FEM simulation responses for the low and middle frequencies was found to be beneficial when the source material contained excessive low-frequency content, or when diffracted waves were prominently present in the BRIR, such as diffracted waves from the car seats.

Postma and Katz [93] conducted a comparison study, where measured and simulated BRIRs of two churches and a theater were compared. GA methods were used in the simulation. Participants were asked to rate the similarity of measured and simulated responses in terms of eight perceptual attributes. For each corresponding responses A and B, the rating was done on a 101-point scale where -50 indicated “A is more ...” and 50 “B is more...”, and 0 indicated no perceivable difference in terms of a particular attribute. The utilized attributes were reverberance, clarity, distance, tonal balance, coloration, plausibility, apparent source width, and listener envelopment. The simulation models were calibrated using the measured responses. It was concluded that in terms of the attributes, clarity and distance were rated to be different between the measured and simulated responses, but mainly the differences were rated to be small when compared with the stability of the participants responses to repeated conditions.

Mehra et al. [82] carried out an experiment where the navigation performance of the test participants in several VR scenes was compared between two different simulation methods. The comparison was done between a GA framework utilizing the image source method, ray-tracing and diffraction modeling, and a wave-based solver utilizing equivalent source method (ESM) [81]. Two different ESM conditions were used: one that included HRTF processing, and one with monaural reproduction. It was concluded that the participants performed the best in the task when the ESM and HRTF processing were used.

Schissler et al. [110] compared the plausibility of an optimized GA approximation scheme with two reference GA simulations with different computational requirements. The participants were presented video pairs where the audio track of each video corresponded to a simulation condition pair optimized - ref-
Experimental methods for perceptual evaluation of room acoustic simulation

Three different video conditions were used, and hence six pairs in total were evaluated by a participant. The task was to give a rating on an 11-point scale in terms of two questions: “which video has the more realistic audio?” and “how similar is the audio in the videos?”. For the first question, a rating of 1 indicated that from the presented video pair AB, video A had more realistic audio, 6 that the audio in the videos were equally realistic, and 11 that video B had more realistic audio. For the second question, a rating of 1 indicated that the audio was very different between the videos, and 11 that the audio was very similar. The stimuli were presented using binaural synthesis. It was concluded that the optimized approximation scheme was not rated as more realistic than the reference simulations, and that the optimized scheme was not rated as more similar to the computationally more demanding reference than the lower quality reference.

Amengual Garí et al. [3] conducted an experiment in which two different VR sound engines were compared. The participants compared the quality of the engines in six different room models and three different source positions per room condition. Two repetitions of each condition were carried out. The conditions were presented with a head-mounted display and headphones. The task for the participants was to choose which of the rendering engines better fit the expectation about the auditory scene based on the visual representation of the space. Additionally, the participant was asked to rate the confidence in their decision on a 5-point scale, and choose three perceptual attributes from a list of twelve in which the engines differed the most. One of the engines provided a very simple image source solution for a box model fitted inside the scene with a single, broad band absorption coefficient and sampled reverberation tail. The second method used a fully dynamic ray-tracing algorithm taking into account the geometric details of the scene. Interestingly, only the second engine took into account the absorption of air in the calculated responses. Both of the engines used binaural synthesis to auralize the GA solutions. The results indicated that the first engine was preferred across the conditions, with an exception of scenes with high reverberation.

A few authors have studied the usage of the FDTD method for auralization. Cobos et al. [30] conducted an experiment where simulated impulse responses calculated with the digital waveguide mesh (DWM) method at different sampling frequencies 20, 30 and 40 kHz were compared with a discrimination test procedure. The responses were low-pass filtered below 5 kHz. The participants in the test were given a task to identify which of the conditions A or B was identical to a reference condition X, where the reference X was either A or B depending on the trial. The method of the stimuli presentation was not reported. It was concluded that the participants could not discriminate correctly between the two highest sampling frequency conditions.

Southern et al. [119] compared the simulated free field responses of two different propagation directions in a FDTD simulation with the SRL scheme using a sampling frequency of 5000 Hz. A discrimination procedure similar to Cobos was
Experimental methods for perceptual evaluation of room acoustic simulation used. Independent variables in the test were low-pass filter cut-off frequency and the distance between the source and receiver. The method of the stimuli presentation was not reported. The authors suggest that below a normalized cut-off frequency of 0.18, the participants could not discriminate between two propagation directions, which both contained some degree of dispersion error.

Some notions can be drawn from the experiments reviewed here. The proportion of GA methods in the perceptual studies is large. This may be partly due to the fact that GA methods have been considered computationally feasible for auralization over broad bandwidth for a long period of time. A recurring method for measuring perceptual attributes has been direct scaling on an interval scale. Regarding the stimuli presentation in the comparison of measured and simulated BRIRs has been a common choice, as the approach does not involve spatial encoding in the case of GA simulation.

Several independent conclusions were reported on the prominence of the late reverberation as an attribute affecting the similarity rating [145, 39]. In the case that the reverberation was considered similar, other attributes were more prominent [93]. As noted by some of the authors, the mismatch may be due to the model parameters such as the choices of absorption characteristics of surfaces. As such, it can be considered difficult to assess the proportion of errors related to the model parameters and the numerical approximation method.

Interesting results regarding the perceived similarity between measured and simulated responses and the perceived match to a visual stimulus was that a very simple model was perceived as being the most similar in comparison to a measured condition [145] and also matching the expectation of the participant when a visual representation of the space was present [3]. The former result may be connected with the notion that disparity in reverberation has a large effect on the similarity rating, and therefore the lack of reverberation is considered less crucial than the misrepresentation of the reverberation in some cases. The latter result, in the present author’s opinion, is an opening to interesting yet potentially complex problem of measuring the perceptual relevance of accurate acoustic simulation in multimodal virtual environments.

A small number of studies have been conducted regarding perceived auralization quality of solely wave-based room acoustic simulation solutions. The conducted studies considered the behavior of the numerical error related to the simulation. These studies [30, 119] give valuable insights on the perception of the error in band-limited responses, but do not provide evidence of the absolute levels of acceptable error.
4. Summary of contribution

The contribution of the studies included in this work can be divided into two categories: auralization and applications. In the following, a summary of the results is given and future directions for the research in these areas and acoustic simulation in general are discussed.

4.1 Auralization

The main contribution of this work is in the quantification of the absolute thresholds for numerical dispersion error in full audible bandwidth compact explicit FDTD simulation in several different acoustic conditions. The results serve as a lower-bound metrics for dispersion error in auralization applications using the FDTD method.

In Publication I, the dispersion error is measured as a function of distance. Previously, the dispersion error in room acoustic context has commonly been quantified using the phase velocity error percentage, and the cumulative error that it produces with increasing simulation distance has had very little attention. The threshold for dispersion error was measured with three different compact explicit schemes with a fixed phase velocity error percentage of 2% at 20 kHz in the worst-case direction of error. The results indicated that with the chosen phase velocity error, the threshold for simulation distance was 9.5 m averaged over the three different schemes. The maximum group delay that occurs at 20 kHz in the measured distance threshold was 1.8 ms. This value can be considered close to the values that previous studies have reported as a perceptual threshold for group delay error. The group delay value decreases towards low frequencies, and it can be speculated that the maximum value at 20 kHz is an optimistic lower bound due to the frequency-depended sensitivity of hearing.

One of the questions that remained open after the first experiment was the matter of how the absorption of air affects the perception of the numerical dispersion error. In Publication II, the absolute threshold for numerical dispersion error was measured under four distance conditions with absorption of air included and the absorption of air excluded. The results indicated that the
numerical dispersion error becomes inaudible with the used distance conditions when the phase velocity percentage falls below 0.28% and the absorption of air is added to the signal.

In Publication V, the audibility of numerical dispersion error was measured in a very simple room acoustic condition that represents a single early reflection from a planar surface that is placed parallel to the median plane of the listener. Two error conditions were under investigation: numerical dispersion error placed in the direct sound, and numerical dispersion error placed in the early reflection. The results indicated that the presence of the reflection does indeed reduce the audibility of the numerical dispersion error in both cases in comparison to the free field condition. Additionally, a paired comparison experiment for simulated room responses with a varying level of numerical dispersion error was conducted. The error in the simulated responses was controlled by varying the sampling frequency of the simulation. The chosen sampling frequency levels were based on the measured threshold values. The results indicated that the participants were able to identify the different error levels, and therefore additional experiments should be conducted regarding the audibility of numerical dispersion error in different room conditions.

4.2 Applications

In Publication III, a proof-of-concept system using distributed real-time FDTD simulation for auralization in AR and VR application was developed. The concept of using distributed systems is topical, as it can be assumed that the form factors of the near future AR and VR devices accommodate neither the computational nor the energy consumption requirements needed to deliver a physically accurate acoustic representation of the scene in real-time. The results of the experiments regarding the performance of the system indicated that the synchronization of the devices associated with the simulation become the main issue when increasing the resolution of the FDTD simulation.

In Publication IV a method to analyze the propagation of an individual, or a combination of modal components inside a general geometry in time-domain was proposed. The analysis is carried out using time windowed frequency analysis of the numerical time-domain solutions at the spatial region of interest. As an example use case, the approach could be used in the analysis of room acoustic problems where scattering surfaces produce large local deviations at particular frequency regions due to the interference of reflected and diffracted wavefronts. From a design standpoint, such a use case could be the analysis of the seating area configurations in performance spaces. The proposed approach is an example of a possible application that large-scale time-domain simulation can offer for room acoustic design and general analysis of room acoustics.
4.3 Future directions

Several different avenues of future research can be identified regarding FDTD simulation and auralization applications utilizing acoustic simulation in general. Firstly, the results of this thesis can be further extended. The audibility of numerical dispersion error in auralization of large spaces and in band-limited conditions is largely unexplored. Studies concentrating on these aspects would have a practical significance. Further research on improved FDTD schemes that are applicable for general room acoustic problems is, to the present authors knowledge, currently ongoing.

In Publication V, a “brute-force” approach for the spatial encoding was used as the processing was possible to conduct off-line, but it is highly likely that a computationally less intense method could be used to produce similar results. With this in mind, evaluation of different spatial encoding approaches for the auralization of FDTD simulated sound fields, or sound fields simulated with any time-domain solver would most probably be of interest for many practitioners. Such studies could lead to improved spatial encoding schemes and possibly solutions to partially mitigate the dispersion error.

An open-source mesh generation engine for common FDTD mesh topologies producing a combination of a structured grid for the internal region of the geometry and fitted cells at the geometry boundaries while taking possible implications for simulation stability into account would most probably be welcomed by practitioners.

A general evaluation of the importance of accurate representation of the sound field in virtual environments, and which representation the content creators find most compelling in the generation of multimodal experiences is largely an open question.
References


References


Errata

Publication I

In Section IV.A the normalized frequency $f_{max}$ for the CCP scheme is typed erroneously. It is 0.176 instead of 0.178.

Publication II

In Equation (1), $f$ should be described as scaled frequency instead of frequency. This error did affect to the experiment itself, but the effect is considered very small (the maximum differences in attenuation that occur with the 344.0 m distance condition for frequencies 1 kHz, 10 kHz and 20 kHz are 0.03 dB, 0.17 dB and 0.60 dB, respectively).

Publication III

The initialism GPGPU is used differently than in this thesis. In this article, the initialism is used for general purpose graphics processing unit, whereas in this thesis it is used for general-purpose computing on graphics processing units. Both versions appear in the literature, but the latter version can be considered to be more appropriate.
Errata

**Publication IV**

Expression for $\beta$ after Eq. (4) should be $\beta = (6 - K)/(2\xi)$.

**Publication V**

In Table I, the values indicating the distinction between the front and the back walls should be $x < 3.5 \, m$ and $x \geq 3.5 \, m$, respectively.