

Sound Field Reconstruction in a room through Sparse Recovery and its application in Room Modal Equalization

Présentée le 23 septembre 2021

Faculté des sciences et techniques de l'ingénieur
Laboratoire de traitement des signaux 2
Programme doctoral en génie électrique

pour l'obtention du grade de Docteur ès Sciences

par

Thach PHAM VU

Acceptée sur proposition du jury

Dr J.-M. Vesin, président du jury
Prof. P. Vandergheynst, Dr H. Lissek, directeurs de thèse
Dr Ph. Herzog, rapporteur
Prof. P. Svensson, rapporteur
Prof. R. Fleury, rapporteur

Nobody in life gets exactly
what they thought they were going to get.
But if you work really hard and you're kind,
amazing things will happen.
— Conan O'Brien

To my family...

Acknowledgements

This thesis would not have been possible without the guidance and full support of many great people along the way.

First and foremost, I would like to acknowledge and show my sincere gratitude towards my thesis supervisor, Dr Hervé Lissek, for giving me the chance to pursue this research. Thank you, Hervé, first, for inspiring me to continue on the journey of Acoustics during my master studies. Your lectures and enthusiasm on the subject have given me great motivation and belief in my career choice. I also will never forget what you did for me during my search for internship at the end of my master program. And most importantly, thank you for your trust and your guidance throughout my doctoral studies. For that I have been able to overcome many challenges along the way. Also I would like to thank my thesis supervisor, Prof. Pierre Vanderghenst for welcoming the Acoustic group as a member the LTS2 lab and for providing me with the opportunity to pursue my research. With respect to my work, I would like to thank the Swiss National Science Foundation for providing me with the financial resources to complete this research.

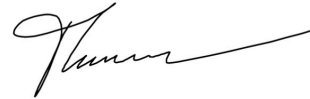
I sincerely thank the members of my jury, Dr. Philippe Herzog, Prof. Romain Fleury and Prof. Peter Svensson who have reviewed my manuscript and provided me with numerous fruitful discussions during my thesis defense. Thank you for all your comments and suggestions for future research directions. A special thanks go to Dr. Jean-Marc Vesin for orchestrating during both my candidacy exam and my thesis defense.

My sincere appreciation goes out to all my colleagues throughout the years at EPFL including members of the past LEMA group, LTS2, LWE and MAG laboratory for all the chat and discussions we have had. My special thanks to all my past and present colleagues at the Acoustic corridor: Baptiste Crettaz, Gilles Courtois, Romain Boulandet, Sami Karkar, Xinxin Guo, Stanislav Sergeev, Maxime Volery who have helped me in various ways during my time in the lab. A warmest thanks to my friends/colleagues Vincent Grimaldi and Thomas Laurence for all the good times and great music room sessions that we had. I wish both of you great luck during your journeys. I'd like to give special thanks to Etienne Rivet for being a great mentor when I first started at the lab. Thank you for always setting an example of a true researcher and a kind colleague.

Acknowledgements

My warmest thanks and appreciation goes to Prof. Kim Yang Hann for all your lectures and guidance 10 years ago. Your passion, determination and greatness have guided me towards this wonderful field of Acoustics. My special thanks also goes to Prof. Cho Ailee for being the one that I could always turn to during my Bachelor years in South Korea.

I would like to express my most sincere thanks to my family back in Vietnam. I would not be the person I am today without my dear parents. Thank you Mom and Dad for bringing me into this world, for your sacrifices and for your unwavering belief in your only child. Con yêu bố mẹ rất nhiều và mong bố mẹ luôn mãi bên con. Lastly, this journey of mine would not have been so enjoyable and precious without the love of my life, my best friend, my wife, Chi. Thank you for always being there for me through thick and thin, for believing in me and most important of all, for your unconditional love.



Lausanne, July 2021

Pham Vu Thach

Abstract

The acoustics of a room often refer to the quality of the room in terms of sound recording as well as sound reproduction. At low frequencies, where the wavelengths are of the same order of magnitude with the dimensions of the room, due to the standing waves phenomena, room modes are formed which can give rise to multiple irregularities and eventually damage the listening experience. In order to effectively address the problem at these frequencies, a thorough knowledge of the sound field in the room is often required. Through numerical studies, the sound field in the room could be simulated based on an FEM room model which would result in certain differences when comparing to the actual situation. On the other hand, the traditional experimental sampling of the sound field in a room often results in an impractical number of microphones.

This thesis focuses on the development of a sound field reconstruction framework that makes use of certain sparse representation and recovery algorithm to reconstruct the sound field of the room at low frequencies using a limited set of microphones. The reconstruction framework is then tested under different conditions and scenarios to prove its validity. Once the performance of the framework has been validated through both simulation and actual measurements, the reconstruction framework is investigated in detail regarding its application in the field of active modal equalization. These active absorption methods are of high interest in the field of modal equalization due to the fact that passive absorption method at low frequencies are more often than not ineffective. The reconstruction framework is then used to first assess the spatial performance of the active electroacoustic absorbers which consists of a current-driven system of loudspeakers in enclosures that aims to dampen room modes by achieving optimal target impedance in front of the loudspeaker membranes. In addition, a more complex applications of the reconstruction framework in which its reconstructed sound field can be used to fine tune the control parameters of the active absorbers is also investigated. The research shows promising results in both application of the reconstruction frameworks and open doors to further development in both the field of sound field reconstruction as well as room modes equalization.

Key words: Sound field reconstruction, sparsity, sparse, matching pursuit, target impedance, equalization, room acoustics, active absorption, modal analysis, mode shape functions

Contents

| | |
|--------------------------------------------------------------------------|------------|
| Acknowledgements | i |
| Abstract | iii |
| List of figures | ix |
| List of tables | xv |
| 1 Introduction | 1 |
| 2 Room acoustics at low frequencies | 7 |
| 2.1 Modal behavior of sound field in enclosure | 7 |
| 2.2 Characterization of room acoustics | 13 |
| 2.2.1 Numerical Studies | 13 |
| 2.2.2 Experimental approaches | 14 |
| 2.2.3 Spatial measurements | 16 |
| 2.3 Perception of room modes and corrective measures | 18 |
| 2.3.1 Room modes and the irregularities of the perceived sound | 18 |
| 2.3.2 Passive and active correction | 20 |
| 2.4 Conclusion | 21 |
| 3 Sparse representation and Sparse recovery | 23 |
| 3.1 Sparsity and Sparse representation | 23 |

| | | |
|----------|---------------------------------------------------------------------------|-----------|
| 3.2 | Matching pursuit | 26 |
| 3.2.1 | Greedy algorithm | 27 |
| 3.2.2 | Matching pursuit | 29 |
| 3.2.3 | Orthogonal Matching Pursuit | 31 |
| 3.3 | Conclusion | 34 |
| 4 | Sound field reconstruction framework | 35 |
| 4.1 | Sparsity in room acoustics at low frequencies | 36 |
| 4.1.1 | Modal decomposition | 36 |
| 4.1.2 | Mode shape approximation | 38 |
| 4.2 | Reconstruction framework | 39 |
| 4.2.1 | Modal identification | 40 |
| 4.2.2 | Projection onto spherical sampled wavevectors | 43 |
| 4.3 | Conclusion | 45 |
| 5 | Reconstruction results | 47 |
| 5.1 | Numerical simulation | 47 |
| 5.1.1 | Modal identification | 49 |
| 5.1.2 | Local interpolation | 51 |
| 5.1.3 | Sound field reconstruction | 53 |
| 5.2 | Experimental results | 63 |
| 5.3 | Conclusion | 67 |
| 6 | Reconstruction of active-controlled sound field | 69 |
| 6.1 | Closed box loudspeaker system and the Electroacoustic absorbers | 69 |
| 6.1.1 | Closed box Electrodynamic loudspeaker system | 70 |
| 6.1.2 | Electroacoustic absorber | 71 |
| 6.1.3 | Existing prototype of electroacoustic absorber | 73 |

CONTENTS

| | | |
|----------|------------------------------------------------------------------------------|------------|
| 6.2 | Sound field assessment of existing active absorption methods | 74 |
| 6.2.1 | Numerical studies | 75 |
| 6.2.2 | Experiment results | 80 |
| 6.3 | Conclusion | 82 |
| 7 | Sound field reconstruction application in room modes equalization | 85 |
| 7.1 | Room modes Equalization metric based on Sound Field Reconstruction | 86 |
| 7.1.1 | Optimal specific acoustic resistance | 86 |
| 7.1.2 | Past optimization attempts | 91 |
| 7.1.3 | Room modes equalization metric using sound field reconstruction . . . | 94 |
| 7.2 | Target Impedance analysis and fine tuning using Sound Field Reconstruction . | 100 |
| 7.2.1 | Prototype | 100 |
| 7.2.2 | Optimal resistance in real experiment | 101 |
| 7.2.3 | Fine-tuning of optimal impedance using Sound field reconstruction . . | 104 |
| 7.3 | Conclusion | 113 |
| 8 | Conclusion and perspectives | 115 |
| A | Additional concepts and formula | 119 |
| B | Additional figures | 121 |
| | Bibliography | 133 |
| | Curriculum Vitae | 135 |

List of Figures

| | | |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1 | Spatial acoustic behaviors of a room mode at low frequencies of the reverberation chamber. | 2 |
| 1.2 | Room Frequency Responses at a single locations in the reverberation chamber. | 3 |
| 2.1 | Standing waves expressed in sound pressure in a 1-D pipe with two closed ends. The resonance frequencies are multiples of the fundamental resonance and depends on the physical length of the pipe. | 8 |
| 2.2 | Typical room modes of a lightly damped rectangular (also called shoe-box) room. | 12 |
| 2.3 | Frequency studies at two different frequencies for a room with complex geometries. | 14 |
| 2.4 | Room Frequency Responses at a few different locations in a room. | 15 |
| 2.5 | The spectrum of the 3D-PAF function visualized by Ajdler in [1, 2]. | 18 |
| 2.6 | Two different microphone placement strategies to increase the accuracy of the sampling using an equi-distant grid of microphones. | 19 |
| 2.7 | The spectrogram analysis of the perceived sound at one location in a reverberation chamber. Due to different modal decay times, some frequencies last longer than others and could results in coloration and masking effects of the perceived sound. | 20 |
| 3.1 | The structure of the formulated Compressed Sensing problem | 26 |
| 3.2 | The activity scheduling problem and its proposed Greedy optimal solution. . . | 28 |
| 4.1 | Modal decomposition/summation for an RFR curve. The response can be seen to be composed of discrete modal responses. | 37 |

| | | |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.2 | Possible star-shaped polygons that satisfy the condition for the recovery of mode shape functions | 39 |
| 4.3 | Examples of room modes in a non-rectangular room. | 40 |
| 4.4 | Example of the spherical sampling that can be used to generate wavevectors for the mode shape approximation process by Paul Leopardi [69]. | 43 |
| 5.1 | Geometry of the FEM model. The black dots represent the measurement points that are spread randomly in the room. | 48 |
| 5.2 | Modal decay times for the eigenfrequencies of the non-rectangular reverberant room as estimated by the SOMP and RFP methods, in comparison with the FEM analysis (reference). | 50 |
| 5.3 | Reconstruction of the RFR at a point inside the room using SOMP and RFP in comparison with the FEM ground truth reference. | 52 |
| 5.4 | Comparison between the RFRs interpolated with RFP (top) and with SOMP (bottom), when underestimating the number of modes, at a given virtual microphone position. The black curve represents the reference given by FEM simulation | 52 |
| 5.5 | A series of 11x11 microphone grids that is used to sample each of the 6 faces of the inner rectangular evaluation area. | 53 |
| 5.6 | Sound field reconstruction (bottom) at different frequencies for a rectangular area inside the room in comparison with the referencing sound fields from numerical simulation (top). | 54 |
| 5.7 | An evaluation 3D grid for the evaluation of the overall accuracy of the reconstructed responses | 57 |
| 5.8 | Sound field reconstruction (bottom) compared to the reference (top) for different cases of wall damping ($\alpha = 0.01, 0.10$ and 0.30 from left to right) at the eigenfrequency around 35.3 Hz | 58 |
| 5.9 | Comparison of the normalized error between different cases of room absorption for the reconstruction of the rectangular volume and a plane near a wall of the room | 59 |
| 5.10 | A randomly redundant set of 600 measurement points in the room that can be used to analyze the dependent of the results on the number of input measurements. | 61 |

LIST OF FIGURES

| | | |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 5.11 | An example of the reconstruction of the RFR at the same evaluation point using two different placements of 15 microphones | 62 |
| 5.12 | Analysis of the Pearson Correlation of the reconstructed sound field with respect to the number of measurement points. | 63 |
| 5.13 | Measurement set up in a real reverberation chamber in the laboratory | 65 |
| 5.14 | Sound field reconstruction from real measurements (bottom) at two distinct eigenmodes (left: around 35 Hz, right: around 51 Hz) as compared to the same ones from simulations (top) | 66 |
| 5.15 | RFR reconstruction from 25 measurements for three different evaluation points in the room using RFP and SOMP (correlation ranging between 97% and 99%). | 67 |
| 6.1 | The control design behind the electroacoustic absorber where the input current of the loudspeaker is controlled by the measured total sound pressure at the diaphragm | 72 |
| 6.2 | The existing Electroacoustic Absorber at EPFL with 4 current-controlled loudspeaker and one input electret microphone | 74 |
| 6.3 | Bode plot of the 3-DOF target impedance in Table 6.2 and its individual 1 dimensional resonators. | 76 |
| 6.4 | Numerical model of the reverberation chamber with the presence of the Electroacoustic Absorbers at 4 corners of the room | 77 |
| 6.5 | Comparing the response of at a single point in space with and without the presence of the 4 Electroacoustic Absorbers | 77 |
| 6.6 | The sound field before (left) and after (right) applying the electroacoustic absorbers are depicted precisely using the reconstruction algorithm (bottom) when compared to the simulation results (top) at $f=45.25$ Hz | 78 |
| 6.7 | The sound field before (left) and after (right) applying the electroacoustic absorbers are depicted precisely using the reconstruction algorithm (bottom) when compared to the simulation results (top) at $f=55.07$ Hz | 79 |
| 6.8 | Analysis of number of microphones vs correlation value | 80 |
| 6.9 | Overall experiment set-up for sound field reconstruction with the Electroacoustic Absorbers. | 81 |

| | | |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 6.10 | The comparison of the reconstructed sound field using SOMP at the first mode (top) when the Electroacoustic Absorbers are turned off (left) and on (right) as compared to the equivalent situation simulated by the FEM model | 82 |
| 6.11 | The comparison of the reconstructed sound field using RFP at the 9th mode of the room (top) when the Electroacoustic Absorbers are turned off (left) and on (right) as compared to the equivalent situation simulated by the FEM model . . | 83 |
| 7.1 | The model of the duct under study ($S_{duct} = 900cm^2$, $L_{duct} = 1.7m$) and the circular absorber with a variable radius and specific acoustic impedance at one end of the duct. | 87 |
| 7.2 | Modal decay time analysis between different combination of the effective absorption area and the purely resistive impedance of the absorber at one end of the rectangular duct. | 88 |
| 7.3 | The model of the eigenfrequency study for the reverberation chamber for variables parameters, namely, the radius of the 16 highlighted speakers of the four Electroacoustic absorbers and their specific acoustics impedance. | 89 |
| 7.4 | Color mapping of the modal decay time of several modes in the FEM model of the reverberation chamber with respect to the specific acoustic impedance and the total effective absorption area | 90 |
| 7.5 | Plot produced by E. Rivet consisting of optimal ranges for the value of specific acoustic impedance for multiple modes of three different room models. The ranges are plotted with respect to the modal frequencies, so as to later form a bounded optimal region for the target impedance curve in the frequency domain. | 92 |
| 7.6 | The 1, 2 and 3 degree-of-freedom settings of the impedance which share the same asymptotes but a wider bandwidth can be achieved using the higher degree-of-freedom models. | 93 |
| 7.7 | The 216 points of measurement grid used for the sweeping parameters study in the numerical simulation. | 95 |
| 7.8 | Average standard deviation of the Room frequency responses at the 6x6x6 sample points with respect to the total effective absorption area and the specific acoustic resistance at the 16 speaker membrane | 96 |
| 7.9 | A comparison between two different responses with different average standard deviation values which shows that the standard deviation could be a misleading metrics in analyzing the equalization aspect of room response. | 96 |

LIST OF FIGURES

| | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 7.10 Comparison of different averaging methods with respect to the original response curve | 97 |
| 7.11 Evaluating the sweep parameters study using two different metrics: a) Average relative moving standard deviation b) Average Root Mean Square Deviation (RMSD). | 98 |
| 7.12 Frequency response (both absolute value and dB value) at a single point in the room under different combination of the sweep parameter study as compared to their location on the color mapping of equalization. | 99 |
| 7.13 The new prototype for the Electroacoustic Absorber | 102 |
| 7.14 The two targeted modes of the optimal resistance test. | 102 |
| 7.15 Optimal specific acoustic impedance test in real experiments for the two first modes of the reverberation chamber. | 103 |
| 7.16 Experiment set-up with 2 Electroacoustic Absorbers at the corners and 15 randomly placed microphones in the room. | 106 |
| 7.17 The 30 stable scenarios for the Target Impedance of the electroacoustic absorbers.107 | |
| 7.18 The equalizing performance of all 30 stable cases of the target impedance in studies. | 108 |
| 7.19 The effects of the center frequency on different settings of the control. Connected points represent settings with the same R_{st} and BW with varying f_c | 109 |
| 7.20 The effects of the bandwidth on different settings of the control. Connected points represent settings with the same R_{st} and f_c with varying BW. | 110 |
| 7.21 Evaluation of metric M2 with respect to R_{st} and bandwidth at different center frequency of target impedance | 111 |
| 7.22 Individual response comparison between two different settings of the controller for target impedance in real setup as compared to the room response without any active control. | 112 |
| B.1 The effects of the center frequency on different settings of the control according to metric M1 | 121 |
| B.2 The effects of the bandwidth on different settings of the control according to metric M1 | 122 |
| B.3 Evaluation of metric M1 with respect to R_{st} and bandwidth at different center frequency of target impedance | 123 |

List of Tables

| | | |
|-----|------------------------------------------------------------------------------------------------------------------------------|-----|
| 6.1 | Small signal parameters of the closed-box loudspeaker system in the existing Electroacoustic Absorber at EPFL | 73 |
| 6.2 | Current target impedance for the 3-DOF resonator model of the electroacoustic absorbers | 75 |
| 6.3 | Correlation value evaluation for the reconstruction from real measurements . . | 84 |
| 7.1 | Small signal parameters of the closed-box loudspeaker system for the new prototype of the Electroacoustic Absorber | 101 |
| 7.2 | Combination of design parameters that results in instability for the control . . | 107 |

1 Introduction

Context

Room acoustics is the domain of study revolving around the physics of sound and how it behaves when enclosed inside the boundary of a room. Due to the existence of the walls, floor and ceiling of a room, the acoustic behaviors of sound in a room are significantly different from its free field behaviors. These behaviors create a diverse range of phenomena that more often than not, damage what we call the acoustic quality of a room [109, 61]. These behaviors furthermore vary depending on the respective wavelength of the sound with respect to the particular room geometries.

At high frequencies, the sound in the room is better understood under the concept of geometrical acoustics or ray acoustics. Within this range, sound behaves not that different from light rays where its reflections can be understood using the image source models. In this range of frequency, problem arising from early reflections and reverberation time could be treated using several placements of passive absorbers and diffusers in the room to absorb sound energy [65].

At low frequencies, where the wavelengths are of the same order of magnitude as the sizes of the room, the sound field is heavily dictated by the presence of the eigenfrequencies of the room, which lead to resonances known as room modes. These room modes are created by the standing waves phenomena that spatially divides the room into regions of nodes and anti-nodes depending on the physics of the particular mode. Due to this phenomenon, the listening experience in the room could be damaged in all three domains of space, time and frequency. Each of these room modes possesses a respective mode shape function that dictates how the sound energy will be distributed spatially in the room. This means that the listening experience in the room vastly differs depending on the location of the listeners as well as the location of the sound source (Fig.1.1). Furthermore, the frequency response (which directly affects the sound reproduction quality frequency-wise) at any individual location in the room will also be characterized by discrete room modes which create the peaks and dips across the entire frequency band (Fig.1.2). This introduces complex coloration and frequency masking

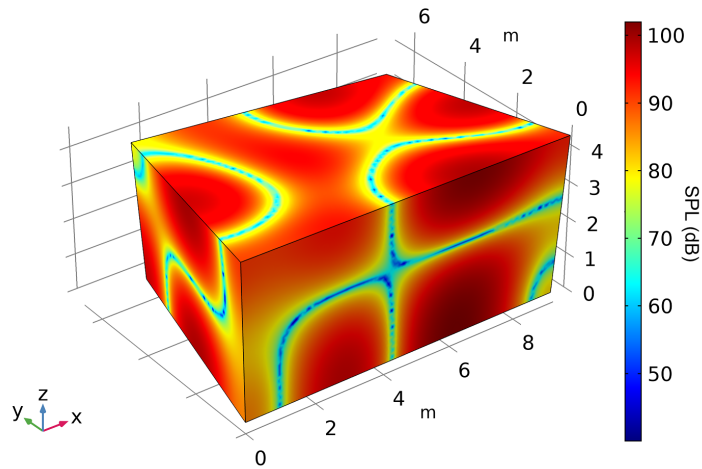


Figure 1.1 – Spatial acoustic behaviors of a room mode at low frequencies of the reverberation chamber.

onto the original sound produced by the source. Lastly, in the time domain, each room mode will possess a different decay times which means that there is an imbalance of decay times depending on the sound emitted by the source. This can create temporal masking effects when a non-stationary sound is played by the source. Due to these reason, without proper treatments, the listening experience at low frequencies in a room will always be heavily colored and distorted by the presence of room modes. At these frequencies, passive absorbers usually cannot effectively damp these room modes, which is why research in recent years have put more focus on designing active absorption methods to equalize the frequency responses of the room.

Motivation

As mentioned previously, passive absorption most of the time finds it challenging to tackle the room modes at low frequencies due to their constraints in sizes. This is why a proper treatment of room modes, more often than not, requires specific knowledge about the modal behaviors of the room.

These knowledge could either be acquired through detailed simulation or practical measurements in the room. The former often makes use of 3D simulation process such as the Finite Element Method (FEM) or the Boundary Element Method (BEM) to either find the analytical eigenfrequencies and their corresponding mode shapes of a simulated room model or to directly simulate the frequency responses in the room resulting from signals emitted by a simulated source. This is often proved to be useful for the understanding of the overall acoustics of the room. However, being a numerical simulation based on mathematical models, this method could not avoid certain inevitable drawbacks. One of the most pronounced drawbacks

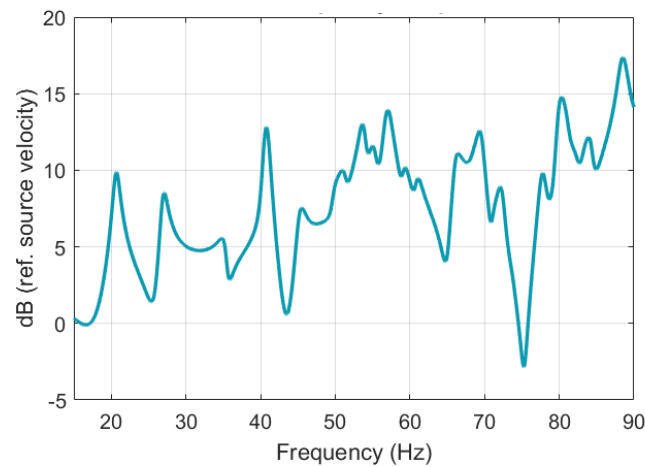


Figure 1.2 – Room Frequency Responses at a single locations in the reverberation chamber.

is based on the fact that it is extremely difficult and laborious to recreate a model close to the actual situation. In terms of room geometries, if there are furniture in the room, this process certainly becomes more complicated. Furthermore, higher complexities arise when we take into account the impedance and absorption characteristics of different materials and furniture shapes which are proven to be difficult to replicate.

Due to these reasons, real experiments are most of the time necessary to analyze the modal behavior of the room. Conventional methods often involve the measurement of the Room Frequency Responses (also known as the Room Transfer Functions) at a few dispersed locations in the room. Through these measurements, the modal parameters of the room could be estimated. However, as the frequency response in the room differs from point to point, a correct modal analysis also requires heavy efforts especially when not much information of the room can be readily available. Another useful information that could be acquired from the frequency responses is the irregularities of room modes across the frequency band and how the peaks and dips of each modes affect the “flatness” of the responses. Once again, this could not be generalized due to the fact that at each frequency, the sound energy in the room is distributed differently due to the existence of room modes.

All of the complications above are caused by the fact that the spatial behaviors of the room are so far unknown. As explained previously, room modes affect the listening experience in the room in all three domains. While the time and frequency domain can be interchangeable, the same cannot be said about the spatial aspect of the responses in the room. This is why the subject of spatial sampling in room acoustics is always an intriguing one but could also often be overlooked by acoustic practitioners due to its intrinsic complication. Conventional sampling of the sound field in enclosures using a regular and equi-distant microphone grids has been shown to require a large number of microphones which often goes into the hundreds to acquire high resolution and accuracy. Nevertheless, the results from these sampling often show that there are aspects of the real sound field that numerical studies like FEM would find

it difficult to capture. Furthermore, studies [2, 99] have shown that a thorough understanding of the sound field of the room may shed new light into how it can be incorporated to benefit active modal equalization.

Objectives

The main objective of the thesis is to develop a robust sound field reconstruction framework that can recover the low frequencies sound field of the room using a much less number of microphones without affecting the accuracy as well as resolution of the reconstruction. We aim to achieve this objective by exploiting the inherent sparsity of the acoustic behavior at low frequency in a room, where each response can be decomposed into a sum of individual discrete room modal responses. Using a sparse representation and recovery theme, we hope to achieve a reconstruction of the spatial responses of the room where the mode shape behaviors could be clearly observed and analyzed with high accuracy.

A secondary objective of the thesis comes from the fact that recent active room equalization methods at low frequencies have called for the development of a method to recover the spatial sound field of the room in order to assist future active control strategies. After testing the validity of our sound field reconstruction framework, we hope to shed light into this matter by experimenting with different ways that a sound field reconstruction algorithm could assess and assist modern active room equalization methods, in particular the Electroacoustic Absorbers developed at EPFL [99]. Through the analysis, we aim to open doors to a wide application of sound field reconstruction in the domain of room modes equalization.

The entire thesis, in overall, combines the knowledge in three different major field of studies, including room acoustics, sparse recovery and active sound control. While the fields of room acoustics and sparse recovery are more intertwined regarding the design and optimization of the reconstruction framework, the applicability aspect of the sound reconstruction framework does incorporate certain aspects of active absorption method using impedance control.

Outline and original contributions

This section summarizes the contents along with the contributions of each chapter of the thesis to different fields of studies.

Chapter 2: Room acoustics at low frequencies

Description: A brief review of physics of sound field in enclosures is covered, from the simplest case of a duct in 1D to a 3D room model. Different methods to characterize room acoustics at low frequencies, from simulation to real measurements are then analyzed. Finally, a brief review on the damaging effects of room modes on listening experience as well as available passive and active methods to dampen room modes are covered.

Contents: The main focus of the chapter is put on the process of characterizing sound field at low frequencies and why it has not been regularly and extensively performed for a room. Different analysis was done on multiple aspects regarding the experimental sampling of sound field to suggest that the existence of a sound field reconstruction framework could be beneficial to multidisciplinary research fields.

Chapter 3: Sparse representation and Sparse recovery

Description: The concept of sparsity and sparse recovery is explained in details to highlights the problem development as well as possible solving strategies for a sparse system or signal. Different approaches are covered from compressed sensing to matching pursuit with the same objective of solving the sparse recovery problem.

Contents: The analysis in this chapter put high focus on the family of Greedy algorithm which includes matching pursuit and briefly discussed its modified versions. From the analysis of room modes in the previous chapter, we highlight the fact that orthogonal matching pursuit could be well suited for the recovery of the sound field of the room.

Chapter 4: Sound field reconstruction framework

Description: Based on the studies from the previous chapters, a sparse representation of the spatial acoustic response of the room can be formed. Through multiple derivations and approximation, a closed form solution for the spatial response can be formed. Finally a framework for the sound field reconstruction from a limited set of measurements is developed.

Original contribution: In this chapter, taking advantage of the sparsity that inherently exists in the wave solution, we develop a reconstruction framework that makes use of multiple algorithm including the Simultaneous Orthogonal Matching Pursuit to recover the governing parameters of the closed form solution of the spatial response.

Chapter 5: Reconstruction results

Description: From the creation of the reconstruction framework in the previous chapter, this chapter aims to put the framework into work using multiple evaluations both in numerical studies and real experiments.

Original contribution: In order to validate the framework from the previous section, we design different analysis and evaluation to test the framework under different condition using the numerical model of a non-rectangular reverberation chamber model. The framework is then proven to perform at high accuracy not only in simulation but also in real experiments where the sound field of the room can be reconstructed from a limited set of measurements.

Chapter 6: Reconstruction of active-controlled sound field

Description: After the successful validation in the previous chapter, in this chapter the reconstruction framework is further tested for situation where there are active absorbers present in the room. The analysis is once again done in both simulation and real measurements.

Original contribution: In this chapter, we proposed to use the reconstruction framework to analyze the performance of active absorbers. In order to test the validity of the framework in this case, the sound field with and without the presence of the electroacoustic absorbers are reconstructed and compare in details to the ground truth results simulated by FEM studies. The study is then extended to the real room where the framework can be used to assess the performance of actual electroacoustic absorbers in the reverberation chamber. All results show a highly accurate reconstruction of the enclosed sound field.

Chapter 7: Sound field reconstruction application in room modes equalization

Description: In this chapter, the application of the reconstruction framework is extended to a higher level where its reconstruction results could be used to assist the design of active absorbers to equalize room modes, such as the electroacoustic absorbers.

Original contribution: In this chapter, we perform some analysis on the design of the electroacoustic absorbers which is based on the target impedance concept of the modal behavior in the room. We then suggest novel equalization metrics that can be gathered from a large set of spatial responses in the room which could be reconstructed using a few measurements thanks to our reconstruction framework. Finally, we experiments with the idea of implementing these metrics to tune the design of the single resonator model of the electroacoustic absorber in which its control parameters can be calibrated based on the performance evaluated using the equalization metrics developed earlier. At the end, the experiments show promising results where not only the parameters of the control can be tuned to achieve high equalization performance for a bandwidth of interest, but also some valuable information of the relationship between the target impedance and the modes of the actual room can be seen through different graphic interpretations.

2 Room acoustics at low frequencies

In this first chapter, we briefly review several aspects regarding room acoustics at low frequencies with the focus on the modal behavior that exists when sound waves are bounded by enclosures. Due to the modal behavior, room modes are formed which results in multiple irregularities that will ultimately damage the listening experience in a room at low frequencies. We also introduce several methods to predict as well as to measure sound field at low frequencies and discuss their strengths and weaknesses. Finally, we briefly review a few existing passive and active methods in room modal equalization, where the objective is to equalize the Room Frequency Responses. This chapter, in overall, serves as an introduction to sound field in a room at low frequencies which builds the foundation for the chapters that follow.

2.1 Modal behavior of sound field in enclosure

The acoustic behaviors of sound in enclosed space are essentially different than in free field. Sound waves that propagates inside an enclosure are reflected by the enclosure boundaries. The characteristics of the reflected wave are altered by the reflection properties of these boundaries. The resulting sound pressure at any point in an enclosed room is a complex combination of the direct sound and its endless reflections off the walls of the room. This creates the phenomenon of modal resonances at low frequencies where the sound pressure will be different at different locations in the room. Furthermore, the perceived sound at a fixed location in the room also vary depending on frequencies. These modal resonances, or room modes and the resulting variations they create in a room is a function of the geometry as well as the properties of the room.

For the most fundamental 1-D case of an air-filled rigid pipe with two closed end where it is assumed that sound waves are not bothered by the other surfaces other than the two closed ends, the pipe can be considered as a 1-D resonator that can resonate at certain characteristic frequencies when excited. We also assume that the length of the pipe is much larger than its diameter so that the sound can be studied purely in 1-D.

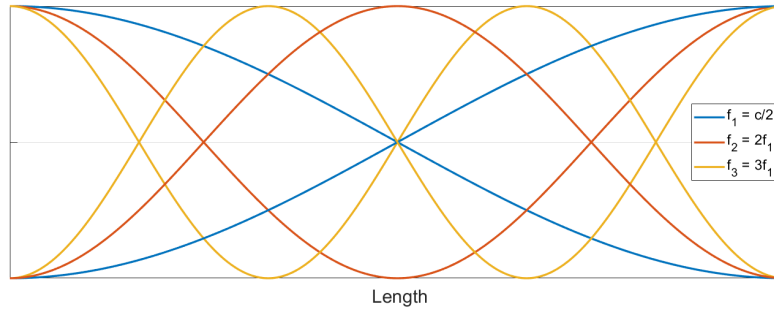


Figure 2.1 – Standing waves expressed in sound pressure in a 1-D pipe with two closed ends. The resonance frequencies are multiples of the fundamental resonance and depends on the physical length of the pipe.

The air in the pipe is then disturbed by an acoustic source somewhere inside the duct at a certain frequency. It is then noticed that there exists certain frequencies where the sound in the pipe is reinforced. The lowest frequency that has this behavior is the first resonance of the pipe f_1 . This phenomenon can also be encountered at $f = 2f_1$, $f = 3f_1$ and other multiples of f_1 . This phenomenon is caused by the interaction between the left and right going waves in the pipe which at resonance frequencies, superimpose to form the standing wave phenomenon. Standing waves are characterized by the locations of its nodes where air particles barely vibrate and anti-nodes where air particles can vibrate with the highest amplitude. These resonant frequencies in the pipe are determined by the length of the pipe by:

$$f_n = \frac{nc}{2L} \quad (2.1)$$

where c is the speed of sound in the air, L is the length of the pipe and n is the index of the resonance. Knowing that frequency can be expressed by its wavelength as $f = c/\lambda$, equation 2.1 becomes:

$$\frac{\lambda_n}{2} = \frac{L}{n} \quad (2.2)$$

which illustrates the modal behavior in figure 2.1 of the first few modes of the pipe. The boundary condition of the two closed end means that pressure will always be maximum at the two ends. Depending on the order of the modes, one or multiple nodes can be found in the space between the two ends of the pipe. This creates a spatial imbalance of pressure inside the pipe.

The previous problem can be extended to the 3-D case, where sound can be studied in an

enclosed 3-D room. There are different methods to approach an acoustic problem in the room. At high frequencies, where the wavelength is several order smaller than the room geometries, a ray-based approach might be applied. At low frequencies, where the wavelengths are of the same order of magnitude as the dimensions of the room, a wave theory approach is more preferred. The starting point of the wave theory representation of the sound field in a room is the Helmholtz equation:

$$\Delta p + k^2 p = 0 \quad (2.3)$$

in which $k = \frac{\omega}{c}$ is the wavenumber. One important assumption that should be made is that the boundaries of the room has impedance that is independent on the direction of incident of sound. In order to take into account the boundary condition of the room, we further have:

$$Z_w \frac{\partial p}{\partial n_w} + i\omega\rho p = 0 \quad (2.4)$$

where Z_w is the specific impedance of the wall, and $\frac{\partial}{\partial n_w}$ indicates a partial differentiation for the direction outward normal to the wall.

It has been shown [65] that the solution to the wave equation satisfying the boundary condition in Eqn. 2.4 is non-zero only for discrete values of the wave number $k = \omega/c$. Calling these values, k_n , we can say that $-k_n^2$ can be considered as the eigenvalues of the wave equation.

When solving using eigenvalues, each of those eigenvalue k_n will be paired with a corresponding eigenfunction. In the case of room acoustic, these eigenfunctions, denoted $\Phi_n(\mathbf{x})$, are called the mode shape functions of the room with \mathbf{x} as the positional vector in the 3-D space of the room. This mode shape function defines the spatial modal behavior of each resonance of the room. These eigenfunctions are mutually orthogonal, which is equivalent to:

$$\iiint_V \Phi_n(\mathbf{x}) \Phi_m(\mathbf{x}) dV = \delta_{m,n} K_n \quad (2.5)$$

with K_n as a constant, V as the enclosed volume in the room and $\delta_{m,n}$ is the Kronecker delta which are always zero and only equal to 1 when $m = n$. The same can be applied to the forced

waves case, i.e, when there is a source in the room. This turns the Helmholtz equation into:

$$\Delta p + k^2 p = -j\omega\rho q(\mathbf{x}) \quad (2.6)$$

in which q is the volume flow of the source. As the mode shape functions form a complete set of the solution, the source could also undergo such expansion:

$$q(\mathbf{x}) = \sum_n Q_n(\mathbf{x}) \Phi_n \quad \text{with} \quad Q_n = \frac{1}{K_n} \iiint_V \Phi_n(\mathbf{x}) q(\mathbf{x}) dV \quad (2.7)$$

Needless to say, this is also applicable for the overall solution to the forced vibration problem:

$$\Phi_\omega(\mathbf{x}) = \sum_n P_n \Phi_n(\mathbf{x}) \quad (2.8)$$

Inserting Eq. 2.7 and Eq. 2.8 into Eq. 2.6 we get:

$$\sum_n P_n (\Delta \Phi_n + k^2 \Phi_n) = i\omega\rho \sum_n Q_n \Phi_n \quad (2.9)$$

For the simple case under study, we assume to have a point source at location \mathbf{x}_S with a source volume flow of q_S . The source function in space can then be represented by a delta function:

$$q(\mathbf{x}) = q_S \delta(\mathbf{x} - \mathbf{x}_S) \quad (2.10)$$

Solving for P_n using Eq. 2.10 and Eq. 2.9 we have the final form of the solution as:

$$p_\omega(\mathbf{x}) = j q_S \omega \rho \sum_n \frac{\Phi_n(\mathbf{x}) \Phi_n(\mathbf{x}_S)}{K_n (k^2 - k_n^2)} \quad (2.11)$$

Due to the complex boundary equations regarding damping of the walls of the room, the complex wave number k_n can be presented using the angular modal frequency ω_n and the

modal decay δ_n :

$$k_n = \frac{\omega_n}{c} + j \frac{\delta_n}{c} \quad (2.12)$$

which leads to the derivation of the solution of:

$$p_\omega(\mathbf{x}) = \rho c^2 \omega q_s \sum_n \frac{\Phi_n(\mathbf{x}) \Phi_n(\mathbf{x}_s)}{K_n [2\delta_n \omega_n + j(\omega^2 - \omega_n^2)]} \quad (2.13)$$

Under the frequency domain, the above function serves as the transfer function of the room between two points \mathbf{x} and \mathbf{x}_0 of the room.

Now we move on to a more specific case of a rectangular room with its side spanning each of the three Cartesian coordinates of x , y and z . the Helmholtz equation can be written as:

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} + k^2 p = 0 \quad (2.14)$$

Following the same procedure for an arbitrary room, it becomes apparent that the modal frequencies and mode shape functions of an un-damped rectangular room with ideally rigid walls, under certain assumptions, can have a closed form expression. For the resonant frequencies of the room, we have the following:

$$f_n = \frac{\omega_n}{2\pi} = \frac{c}{2} \sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2} \quad (2.15)$$

in which L_x, L_y, L_z are respectively the sizes of the room in the Cartesian coordinates. $n_x, n_y, n_z \geq 0$ are the corresponding indexes of the 3-D mode.

The corresponding mode shape function also has a closed form expression which is composed of the contribution of the axial modal behaviors in each of the three pair of parallel walls (including the ceiling and floor of the room):

$$\Phi_{n_x, n_y, n_z}(x, y, z) = \cos\left(\frac{n_x \pi x}{L_x}\right) \cos\left(\frac{n_y \pi y}{L_y}\right) \cos\left(\frac{n_z \pi z}{L_z}\right). \quad (2.16)$$

Due to this formulation of the mode shape function of the room, the room modes in a rect-

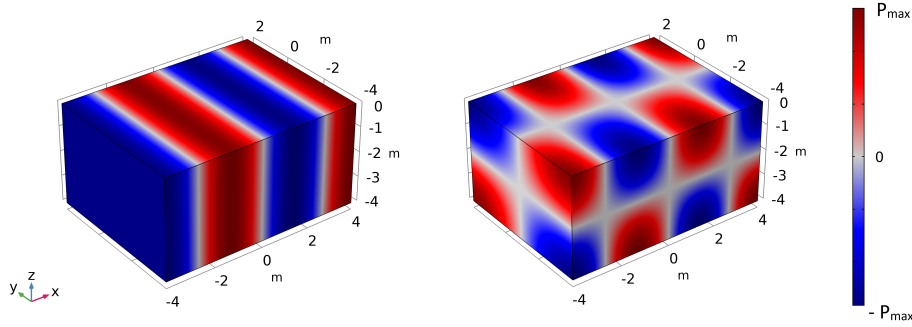


Figure 2.2 – Typical room modes of a lightly damped rectangular (also called shoe-box) room.

angular room can be categorized into axial, tangential or oblique modes. If only 1 out of (n_x, n_y, n_z) is non-zero, the mode is said to be an axial mode, where pressure only fluctuates in 1-D. If 2 out of the 3 are non-zero, the mode is said to be a tangential mode, or a 2-D mode. Finally, if all three of the indexes are non-zero, the mode is said to be oblique which is the most complex type of modes space-wise for a rectangular room. Figure 2.2, showcase two different modes of the rectangular room.

When the walls are non-rigid, however, the analysis of the modal behavior in the room becomes more complex since not just the resonances of the room are shifted frequency-wise, but its mode shape functions are also more difficult to analyzed [83, 13].

From the calculation in Eq. 2.15, we can also see that the number of modes increases rapidly along the frequency axis. This rapid increase in the density of the modes could be estimated [65] by:

$$\frac{dN}{df} = 4\pi V \frac{f^2}{c^3} \quad (2.17)$$

which estimates the density of the eigenfrequencies of the room by a function of frequency. As the frequency gets higher, there are more and more modes contained within a fixed frequency range. When the average half-width of the resonance is much larger than its spacing with neighboring resonances, modes will overlap each other and makes it even harder to detect. Hence, it was stated in [106] that there is a limiting frequency where the wave theory analysis becomes less efficient to analyze the sound field in the room due to the fact that the room starts to act more like a reflector/diffusor model rather than a resonator model. This is called the Schroeder frequency and can be estimated by:

$$f_s = 2000 \sqrt{\frac{T_{60}}{V}} \quad (2.18)$$

where V is the volume of the room and T_{60} is the Sabine's reverberation time [41, 28, 105] of the room which is related to the average absorption of the room.

2.2 Characterization of room acoustics

2.2.1 Numerical Studies

Section 2.1 presented a wave-based simplified approach in analyzing the acoustics of a room at low frequencies. For a rectangular room with rigid walls or very low damping, the analysis of the sound field can result in certain closed form expression of important aspects such as eigenfrequencies and mode shape functions. However, for a given enclosure without these assumption, the explicit analysis of the acoustics of a room at low frequencies is in general, quite difficult. In this case, the use of numerical methods such as Finite Element Method (FEM) are usually involved.

FEM software are very useful and attractive because of its capability to solve complex partial differential equations as well as convoluted interactions between multiple domains. It especially becomes handy in the case for room acoustics due to its ability to handle arbitrary room geometries and surface properties.

The quickest and most simple analysis that FEM can be used for room acoustics at low frequencies is to extract the eigenfrequencies of the room. The results of this would be a list of complex eigenfrequencies of the room. A more sophisticated and computationally heavier analysis is the study of the responses in the room where the acoustics of every points within the mesh of the room is constructed based on the geometries, boundary conditions, as well as the characteristics of the source.

FEM inherently is based on the solution of Helmholtz equation for room acoustics. It makes use of a meshed model where a large number of inter-connected volume elements are linked by analytic nodes and regulated by the boundary conditions of the room. Within the same frequency range of interest, FEM tends to require a considerably higher computational cost than methods based on the geometrical acoustic studies which treat sound under the concept of sound rays. Individual mesh size for an FEM study is usually advised to be at least one-sixth of the shortest wavelength under study (which points to the upper bound of the frequency range of interests). Figure 2.3 presents the Frequency study for a complex room shape with non-zero wall absorption at two different frequencies of 35 Hz and 70 Hz.

The disadvantage of using a numerical analysis, besides its computational load, is the difficulty in building a model close to the actual situation. For an empty room, room geometries seem to be the easiest to replicate in an FEM software as small mismatches does not deviate the results too much, especially when analyzing the room at low frequencies. When there are furniture in the room, however, the study design should take into account the compromise between computational cost and the accuracy of the model. A too detailed model takes away the

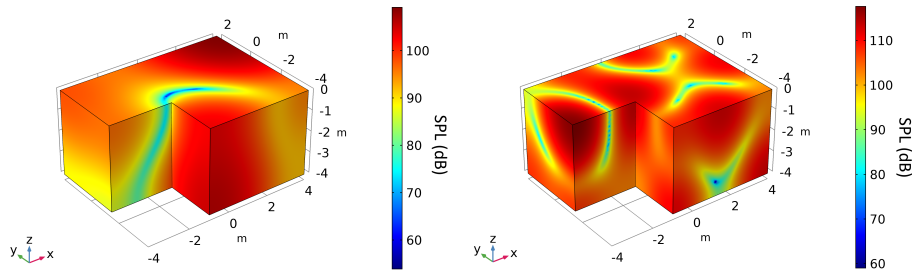


Figure 2.3 – Frequency studies at two different frequencies for a room with complex geometries.

convenience of using FEM, while a relatively too simplified model might result in higher errors. A more complex problem arises when analyzing the properties of the room, more specifically the impedance of the surfaces. Modelling surfaces is one of the most difficult aspect in FEM, especially when we consider the characteristics of the walls as well as the furniture. Without knowing the correct materials and its corresponding acoustic characteristics, it becomes difficult to model them correctly in FEM. Furthermore, even when the exact materials are known, the absorption coefficients, especially at low frequencies (≤ 100 Hz) have been recently shown to be difficult to correctly measured in real life using the traditional analysis of the reverberation time RT_{60} in the reverberation chamber or impedance tube ([96, 58]). Moreover, most walls cannot be described by an impedance because of possible non-local reactions. Due to this reason, FEM should not be used as the only reference when analyzing the acoustic of a room. Instead, it should serve as a powerful tool that can be used to predict and analyze the acoustics in the room on a wide varieties of circumstances and scenarios which might be proven difficult or time consuming if done through real measurements.

2.2.2 Experimental approaches

From the previous section, we can observe the capabilities of numerical studies on assisting the analysis of room acoustics at low frequencies. Due to the complexities of the modelling process, practical measurements are most of the time still necessary to either confirm a theory or to gather the exact actual situation of the acoustic in actual rooms.

In the case of room acoustic at low frequencies, measurement usually refers to the Room impulse response (RIR) or its Fourier equivalent, the Room Frequency Response (RFR) between a source to one or multiple receivers. Individual custom measurements can also be done to extract certain metrics for room acoustics such as the modal decay times and background noises.

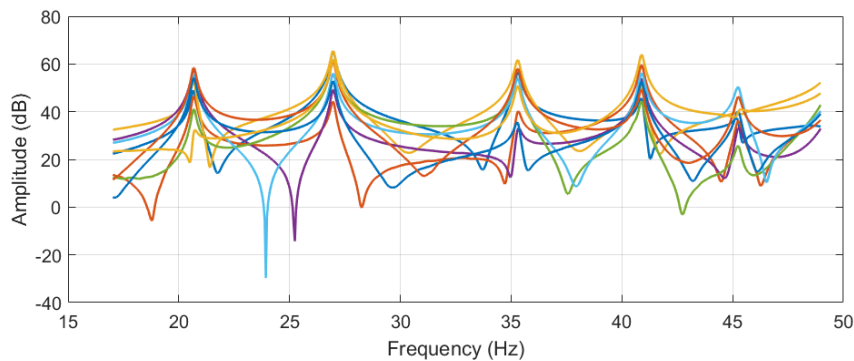


Figure 2.4 – Room Frequency Responses at a few different locations in a room.

Discrete measurements in the room

Measuring sound pressure as a function of frequency is one of the key measurement for room acoustics at low frequencies. It can easily be done by placing a microphone at the location of interest and measure its frequency responses. In order to have the Room Transfer Function (also can be called Room Frequency Response), a reference measurement of the source is also required as an actual source does not have a flat response on the frequency axis. In this case, a velocimeter can be used to measure the velocity of the source as a reference for the Room Frequency Response. A traditional RFR measurement with sweeps [85] or methods using the H1 or H2 estimator can be used to calculate the transfer function between the source and the receivers through the use of cross-spectral and spectral power density [8, 67]. Using these technique, the frequency response at a point in the room can be measured. Depending on the location of measurements, each discrete measurement will possess a different shapes with emphasizes on some of the room modes. In order to capture most of the peaks of the modes of the room, sometimes, measurements at a point close to the corner of the room is preferred. Fig. 2.4 shows the RFR at a few different locations in the same room.

One problem with the Room frequency responses is that a small number of disperse measurements in the room could not tell the whole story regarding how the sound field in the room is spatially distributed. This becomes even more crucial when the room possesses an irregular shape.

Modal decay times measurements

Sometimes, similar to the eigenfrequency studies in FEM, we wish to extract the modal information of the room through real measurements. This process although seems straightforward, is not as simple as it seems. One of the most frequently mentioned metric for this is the Modal decay time (MDT) which describe how long it takes for the decay of sound pressure level at a

mode to reach -60dB. It is linked to the modal decay δ_n through:

$$MT_{60n} = \frac{3 \ln(10)}{\delta_n} \quad (2.19)$$

In order to measure this modal decay time, the simplest way is to use a sine wave generator and target the frequency of a mode. The monotone is generated and then cutoff, the modal decay time of the corresponding mode can then be measured using the decay curve after cutoff using the protocol set by the International Organization of Standardization (ISO 3382-2-2008). Like other types of decay time measurements using the decay curve, certain errors regarding the extrapolation of the decay curve as well as background noises should be predicted. Moreover, additional errors might exist if either the mode has high damping or there are neighboring modes that are too close to the mode of interest. This will, in turn, make the decay curve harder to interpret. Furthermore, doing this require the knowledge of the exact location of a mode and for higher order modes, this is also not a simple task.

A more profound method to estimate the modal decay times is from estimating the damping of a mode using the frequency response curve. Using the bandwidth and quality factor Q of the mode, ones can estimate its modal properties. Due to the fact that sometimes not all the modes can be visibly present on the response curve, multiple methods based on polynomial parametric model such as the global curve fitting using Rational Fraction Polynomial [52, 98, 43] or global Vector Fitting [50, 51] could be used to analyze a series of frequency responses measured at different locations in the room. These method, despite their robustness, always require supervision from users as sometimes incorrect modes are found that might damage the entire estimation.

2.2.3 Spatial measurements

The measurement methods mentioned in the previous sections omit an important features of the modal behaviors of room acoustics at low frequencies, its spatial responses. This is something that a few metrics or response curves could not express. The knowledge on the spatial response of the room could bring a different but equally valuable observation on the acoustics of the room, especially if the room has multiple furniture or has a non-rectangular shape.

In order to perform a spatial response measurements, certain spatial sampling regime should be established. The most important aspect of this measurements is the number of measurement points must be high to avoid both the spatial aliasing and the interpolation errors. In [61], it was suggested that an equi-distant grid with spacing smaller than $\lambda/8$ should be used to fully captures the spatial response inside of an enclosure. In [40], a 100 points grid was proven sufficient to visualized the spatial responses in a small car compartment at 77 Hz. It was then

shown that this spatial analysis usually bring about multiple differences when compared to the FEM model due to the unavoidable simulation difficulties of absorption characteristics.

In [1], a global study under the scope of a signal processing point of view on the sampling of the so-called spatio-temporal plenacoustic function was done by analyzing the spectrum of the function. The study aims to create a global scheme for the uniform spatial sampling of acoustic signal in space and analyze the sparse spectrum of the 4-D Fourier transform (3-D in space, 1-D in time) of the plenacoustic function. The particular research analyzed the mathematical background and relation between spatial resolution and temporal resolution if ones desire to retrieve the spatial responses of the room. For the time domain, the Nyquist-Shannon sampling theorem already stated that the sampling frequency should be set higher than twice the cutoff frequency of the measured signal. Based on this, different criterion for the spatial sampling should also be satisfied.

From the analysis in [2], it was found that for a 1-D sampling on a line, a 2D-FT on the spatio-temporal wave solution can bring out certain aspects of the spectrum of the acoustic function. This was done by taking into account both the temporal angular frequency ω (rad/s) the spatial angular frequency (denoted ϕ_x (rad/m)). Performing a 2D-Ft (1 in time, 1 in space), the 2D-FT function of $\hat{p}(\phi_x, \omega) = FT\{p(t, x)\}$ was found to have the energy of its spectrum concentrated within a triangle bounded by $|\phi_x| \leq |\omega|/c_0$. If ϕ_x then exceeds this bound, the corresponding acoustic function decreases faster than an exponential. Using this as the foundation, a sampling theorem was developed depending on the signal-to-noise criteria SNR:

$$\frac{2\pi}{\delta_x} \geq \frac{2\omega_c}{c} + \epsilon(SNR, \omega_c) \quad (2.20)$$

where ω_c is the cutoff angular frequency. Expanding the study onto the 2 dimensional space, the PAF 3D-FT spectrum is now within the supports of a 3-D cone confined by $\phi_x^2 + \phi_y^2 \leq \omega^2/c_0^2$. Finally, for a 3 dimensional space, the spatial temporal acoustic functions is concentrated within a 4 dimensional hypercone $\phi_x^2 + \phi_y^2 + \phi_z^2 \leq \omega^2/c_0^2$ in Figure 2.5. Using this knowledge on the spectrum of the PAF, the research also suggested different strategies to sample the responses in space. This could be shown to either reduce the number of measurement points or increase the SNR of the reconstruction when interpolating on a line or in 2D using simulations.

There are several difficulties, however, when analyzing the sampling in a confined space, especially for the low frequencies. Practically, using a regular grid to sample an enclosed sound field would prove difficult to achieve high accuracy if it simply follows the limit above and the Nyquist theorem. The reason is, unlike the simple sampling scenario in time, there are physical limits by the boundaries of the confined space. The interpolation for a regular sampled grid theoretically uses a convolution with a sinc function with an infinite support.

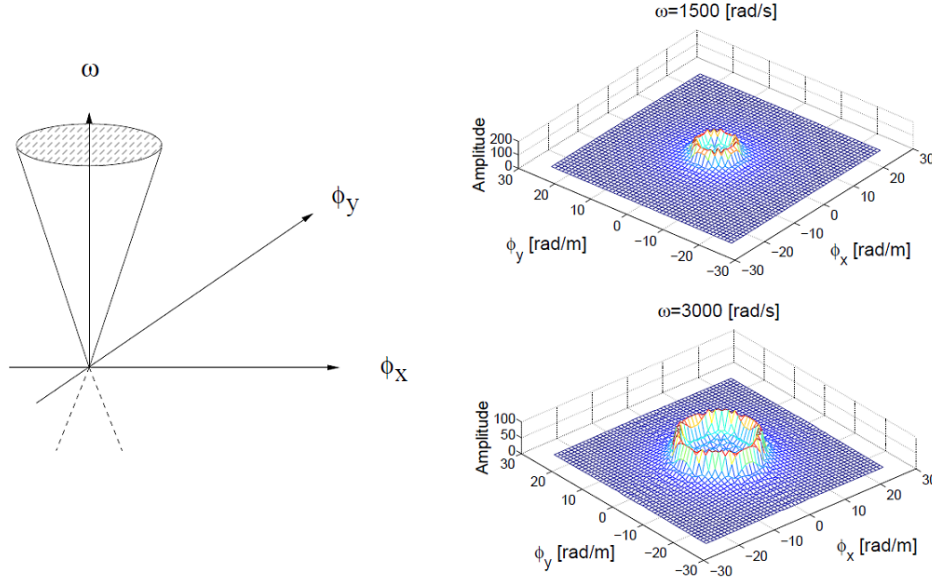


Figure 2.5 – The spectrum of the 3D-PAF function visualized by Ajdler in [1, 2].

This could also become clearer from the Petersen–Middleton theorem [93]. In time, this does not pose too much problem as the extents of both time and frequency of measurements is achievable. In space however, the sampling space is limited and hence for a regular grid sampling, as has been done in the case in [40]. In order to achieve high accuracy, the distance between microphones has been reduced into $\lambda/8$ to increase the resolution within the domain. A second strategies to do this is to keep the sampling rate from Nyquist and [2], but reducing the sub-domain of reconstruction which results in a much larger grid of microphones outside the domain of interest (Fig. 2.6. Both of these strategies, in practice, ultimately still results in a large number of microphones. Later on, in Chapter 4, we will introduce a framework that uses a randomized placement of microphones which can help reduce the number of microphones as well as increase the level of accuracy using a combination of different sparse recovery techniques.

2.3 Perception of room modes and corrective measures

In this section, we discuss the effects that the room has on the perceived sound due to the existence of room modes at low frequencies. Afterwards, several corrective measures, both passive and active can be briefly introduced.

2.3.1 Room modes and the irregularities of the perceived sound

Due to the discrete resonances of the room, the perception of sound in a room at low frequencies undergoes changes that would inevitably damage the listening experience. The

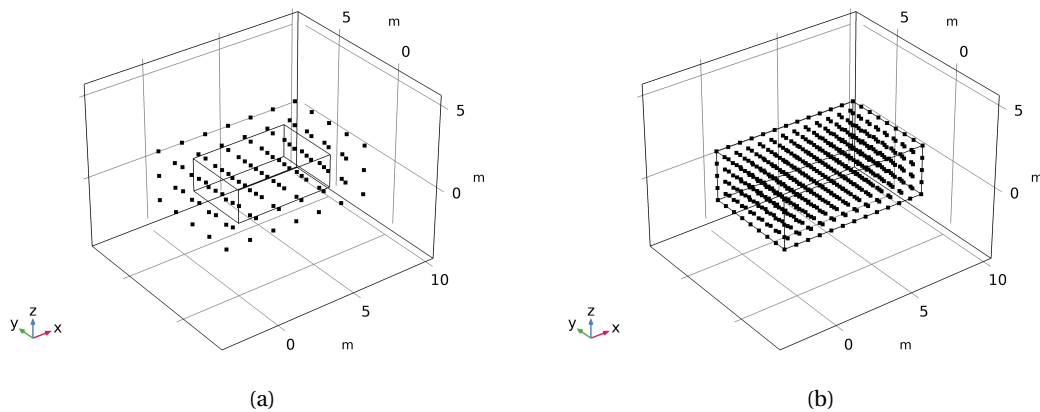


Figure 2.6 – Two different microphone placement strategies to increase the accuracy of the sampling using an equi-distant grid of microphones: a) Keeping the microphone distance but reducing the reconstruction space, b) Keeping the same reconstruction space but increasing the number of microphones (i.e reducing distances between microphones).

psychological aspects of room modes and its irregularity, though not the main focus of the thesis, is an interesting subject nevertheless with various contributions from different authors in the past [42, 95]. These irregularities happen for all three of the domain of interest, i.e time, frequency and spatial domain.

In the frequency domain, it can be seen from the Room Frequency Response of the room at a single location that the energy distribution frequency-wise is not flat. This means that the perceived sound will be significantly colored comparing to the original output of the source.

In the time domain, the existence of room modes and their respective different modal decay times can cause ringing effects in the room as well as coloring and masking effects on the perceived time signal. For instance, when a music piece is being played by the source, certain frequencies of the music piece will have a much longer decay time than other. If, say, a note in the music piece is exactly at one of the eigenfrequencies, this note, besides being amplified as discussed in the previous paragraph, will also have a ringing long decay which might be able to mask some tones from the music piece that comes after it. This, obviously is not what was intended from the original recordings. This means that the perceived sound will be colored as well as will possible lost some of its perceivable notes due to long modal decay times. Moreover, a note played close to these resonances could have its decay shifted towards these eigenfrequencies. Figure 2.7 presents the Short time Fourier Transform (STFT) of a white noise played in the reverberation that when cut off, reveals the effects the room modes have on the perceived sound. The decay tails of the modes of the room can be clearly seen on the spectrogram.

In the spatial domain, it is not difficult to see that due to room modes, the spatial distribution of sound energy is not balance space-wise. Furthermore, this distribution changes as a function

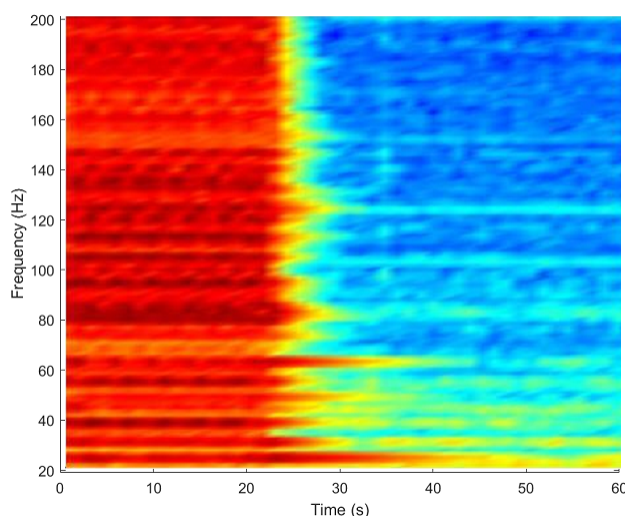


Figure 2.7 – The spectrogram analysis of the perceived sound at one location in a reverberation chamber. Due to different modal decay times, some frequencies last longer than others and could results in coloration and masking effects of the perceived sound.

of frequency. This eventually means that two arbitrary different locations in the room will have a completely different experience of the perceived sound coming from the same source. All the coloring and masking effects previously discussed will also have different interpretations depending on the location in the room. Furthermore, this also creates a vice-versa difficulties when placing the source in the room as the location of the source will heavily affect how it was perceived at different locations in the room.

2.3.2 Passive and active correction

Due to the effects they have on the perceived sound, various attempts, both passive and active have been studied to equalize the responses of the room.

One of the most traditional method to increase the damping of the room is to add more passive foam-based absorbers in the room. However, due to their size dependent on the concerned wavelength, they are not suitable for damping low frequencies below 100 Hz. More advanced method such as porous absorbers could be put in the room to damped room modes. However, as discussed in [61, 5, 64], the ideal location that these absorbers might flourish lies somewhere around the middle of the room where the particle velocity is maximal. Further design also includes Helmholtz resonators [44, 80] but due to the specific design, they could not perform well for a wide range of frequencies.

Another corrective measures is in the domain of architectural acoustic, where the room geometries can be designed to satisfy certain criteria regarding room modes. There is a diverse field of designs which based on many different criteria, such as the Bonello criterion [88].

These designs, although can compensate for a few drawbacks of room modes, technically don't tackle the source of the problem which is the un-damped room modes.

When the room could not be treated using additional absorbing materials, the sound source could be optimized to compensate for the effects of room modes [23, 74, 17]. Furthermore, there also exists methods that borderline on the active side of room mode treatment using the source by means of active equalization of the loudspeaker. These methods are effective to a certain extends but needless to say, will not work if the sources are live instruments.

On the active side of room modes control, different single-input-single-output equalization [76, 6] as well as single-input-multiple-output and multiple-input-multiple-output [54, 81] system could be used to equalize the frequency response function for a single point or multiple points in the room based on multiple principles ranging from active noise cancellation [90] to wave-front separation and echo minimization [53, 77].

An interesting direction that has recently shown highly promising results in equalizing the frequency responses of the room is active control based on impedance control. These methods in general, use the loudspeaker (different from the source) as a controlled absorber where the impedance in front of the membrane can be actively vary to satisfy certain objective functions. These methods have a wide varieties of structure, ranging from a shunt-based control [71, 72] to self sensing method [55, 36, 15]. In Chapter 6, when testing our sound field reconstruction framework with active room control methods, we will review further a method based on the hybrid impedance control behind the Electroacoustic Absorbers at the facility in EPFL.

2.4 Conclusion

In this chapter, we have discussed room acoustics at low frequencies and multiple aspects revolving around the subject. The behavior of sound field in the room at low frequencies which is governed by the standing waves phenomenon has given rise to irregularities in all three domain of time, frequency and space. There are several methods to capture these behaviors in the room. One of the most common method is to use numerical studies through the use of an FEM or BEM analysis. This method, though very useful in giving us the first look at the room's behaviors at low frequencies, could not entirely substitute real measurements as it is to this day, still a challenging task to correctly model the room and its physical properties.

Another route is to perform real measurements in the room. One of the most frequently used method is to measure the Room Frequency Responses at a few locations in the room. However, as the sound field in the room also varies in space, a few scattered measurement could not capture the spatial response of the room at each frequency. This is why in many cases, it is necessary to spatially sample the sound pressure in the room to recreate its spatial responses. However, traditional methods using a regular rectangular grid could result in a large number of microphones, often goes to the hundreds to capture the enclosed sound field with satisfactory resolution. This is one of the main motivation of our research which

is to develop a comprehensive framework to reconstruct the sound field of the room at low frequencies. As the room responses at low frequencies are governed by only a discrete number of modes, we could be inclined to re-model the reconstruction process as a sparse recovery problem. The next chapter will discuss in details this domain of research and suggest how it could benefit a sparse reconstruction framework for the sound field of the room, which will be proposed in Chapter 4.

In this chapter, we have also discussed the damaging effects that room modes have on the listening experience by analyzing the responses of the room in time, frequency and spatial domain. In order to lessen these unwanted effects at low frequencies, different room modes equalization methods, both passive and active have been briefly reviewed. Recent active methods based on impedance control have shown promising results in equalizing the sound field of the room. However, there are still rooms for improvements especially from the experimental point of view. This will be analyzed in more details in Chapter 7.

3 Sparse representation and Sparse recovery

In this chapter, we discuss the theme of sparse representation and sparse recovery. We first discuss the very meaning of sparsity and why it is valuable to the world of signal processing and system identification. Followed by that is a brief review of some powerful algorithms that have been developed to exploit the sparsity that exists in complex system/signals. These include compressed sensing, which is a domain of research that focuses on recovering the most sparse representation of the signal through the use of the optimization of the ℓ_1 norm instead of the traditional ℓ_2 or the problematic ℓ_0 norm. A different approach which makes use of a Greedy Algorithm approach and generates an over-complete dictionaries of possible solutions is the Matching Pursuit algorithm and its multiple modified successor. This chapter serves as a foundation to the use of these sparse recovery regimes in the next chapter where the sound field in a room at low frequencies can be reconstructed from a few dispersed measurements at different locations in the room.

3.1 Sparsity and Sparse representation

The word "sparsity", in its most conventional way, describes the condition of not having enough of something. In the domain of system identification and signal recovery, the description becomes more specific, most importantly focusing on a term that often means nothing: "zero". This section aim to briefly discuss subjects revolving around sparsity - the state of being sparse and why it is of high importance in the world of signal recovery and identification.

Sparsity, or the state of being sparse, can be most easily understood under the frame of a matrix. A sparse matrix is a matrix that consists of mostly zero elements. In contrast to this, a matrix that contains mostly non-zero entries can be defined as a dense matrix. From a system point of view, a full matrix usually describes a complex system with multi-dimensional interactions between its so-called moving parts or state variables. A sparse system, on the other hand, have a more limited number of these interactions and connection.

When data management is involved, sparsity plays an important role in compressing and

storing of information. If a traditional storing or saving process is done on sparse matrices, both processing time as well as memory slots are wasted on a large number of zeros within the matrices. With the knowledge that a matrix is sparse, a more convenient and compressed method can be used to store it without losing its properties. The simplest example that can be analyzed is the case of a diagonal matrix where all its non-zero entries are stored within its main diagonal. Storing each and every entries of this matrix would mean increasing the processing time and memory by an unnecessary amount that keeps growing as the size of the diagonal matrix increases. The most obvious solution to this would be to store the diagonal entries of the matrix in a 1-D form which can reduce the amount of data by a square root of 2.

Previously, we have briefly discussed the importance of sparsity from the point of view of data compression. Conversely, its dual problem can be observed under the scope of system identification and signal recovery/reconstruction. Two of the most important and common questions that ones can run into in this domain is: "If the system is sparse, how much of it is sparse?" and "If the system is sparse, should and would there exist a more efficient way to identify it from a smaller set of sampling data?". These intriguing question becomes clearer if we approach the problem though the form of an under-determined system of linear equations:

$$\mathbf{y} = \Phi \mathbf{x} \quad (3.1)$$

where Φ is an $M \times N$ under-determined matrix with $M < N$, consequently $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{y} \in \mathbb{R}^M$. Matrix Φ is the system dictionary matrix and can be considered as full-rank in this case. The sparse analysis (or sparse approximation problem) can be proposed as follow: Given that the dictionary matrix is under-determined in this case, which inevitably results in an infinite number of solution, find the solution that will have the least number of non-zero entries, i.e, the most sparse solution.

From a mathematical point of view, this can be systematically formulated into an optimization problem:

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^N} \quad & \|\mathbf{x}\|_0 \\ \text{s.t.} \quad & \mathbf{y} = \Phi \mathbf{x} \end{aligned} \quad (3.2)$$

where $\|\mathbf{x}\|_0$ is the ℓ_0 norm of \mathbf{x} which count the number of non-zero components of \mathbf{x} . This problem in general is computationally considered to be NP-hard (as hard as the hardest problem in non-deterministic polynomial time(NP)) considering the use of the ℓ_0 norm. Basically, the problem does not simply ask for any solution but instead ask for the recovery of the most sparse signal.

The equation in 3.2 represents the sparse decomposition process where if \mathbf{x} is sparse, it will only be composed of partial of the matrix Φ . From the signal recovery point of view, \mathbf{y} can be considered as the sampling of size M of the signal(s) \mathbf{x} with size N . Now, if there are at least as many measurements as unknowns ($M \geq N$), assuming that Φ is full-rank, then the problem is exactly determined or over-determined. In this case, there exist multiple ways to find \mathbf{x} (either using the row echelon form through the Gaussian elimination method or through a least square approximation). In the case we are considering, where $M \leq N$, sparsity in \mathbf{x} , if exists, may provide us with a unique solution. However, once again, as explained before, the original problem that optimize using the ℓ_0 norm is NP-hard so there need to be other ways to approach it other than the direct way.

In practice, the measurements from \mathbf{y} are affected by noise. This calls for a relaxation of the equality condition by replacing it with the constraints on the ℓ_2 norm of the error of the solution:

$$\|\mathbf{y} - \Phi\mathbf{x}\|_2^2 < \epsilon^2 \quad (3.3)$$

Presenting the optimization in the Lagrangian form, we have:

$$\min_{\mathbf{x} \in \mathbb{R}^N} \lambda \|\mathbf{x}\|_0 + \frac{1}{2} \|\mathbf{y} - \Phi\mathbf{x}\|_2^2 \quad (3.4)$$

with the Lagrange multiplier λ playing the role of bounding the error of the fitting. This problem, not too different from the one in 3.1, is NP-hard and proved difficult to solve. Two popular but very different approaches to this problem can then be introduced. The first solution comes from the fact that the ℓ_0 norm is replaced by the ℓ_1 norm. This is based on the fact that it was proven in [21, 20] that for a vast majority of problem, solving for the ℓ_1 norm is equivalent to solving the ℓ_0 norm. This opens door to formulate the problem in terms of a compressed sensing framework ([31, 34]) which exploits the sparsity of a signal or system in order to recover signal with a far fewer samples than required by the sampling theorem. Techniques relating to the solving of the ℓ_1 norm includes linear programming ([30, 68, 94]) as well as basis pursuit denoising ([27, 46]). In compressed sensing, it is proposed that based on the sparse basis of the signal, a nonadaptive sampling regime can condense the signal into a significantly smaller amount of data. Assuming that \mathbf{x} is inherently sparse but can only expressed such sparsity under the basis of transformation Ψ such that:

$$\alpha = \Psi^T \mathbf{x} \quad (3.5)$$

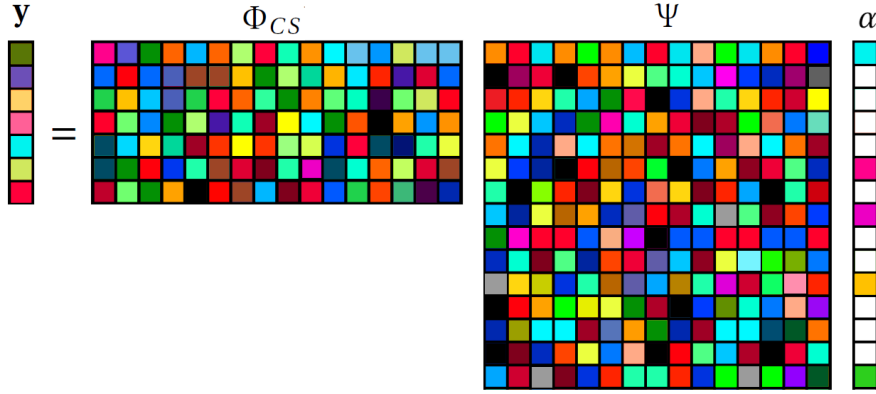


Figure 3.1 – The structure of the formulated Compressed Sensing problem

in which the coefficients contained in α is sparsely distributed. One of the example of this is the wavelet basis [78, 3] that was used to compressed an image. Natural images are dense in the Fourier transform domain but sparse in the wavelet domain. This is due to the fact that wavelet transform possess the time localized property besides its frequency based one. This knowledge leads to the compression of image under the wavelet transform that are presented in the JPEG 2000 format [107, 116, 112]. Coming back to the problem at hand, compressed sensing then poses the sampling governing problem as:

$$\mathbf{y} = \Phi_{CS} \Psi \alpha + \eta \quad (3.6)$$

where Φ_{CS} is the compressed sensing measurement matrix. With the assumption that α is sparse, the incoherence of Φ and Ψ means that the information captured in \mathbf{x} is fully present in the measurements \mathbf{y} . This, as well as the Restricted Isometry Property (RIP) condition in [22, 11, 39], make the problem manageable. The compressed sensing structure is illustrated in Figure 3.1.

The second approach makes use of the greedy algorithm along with a redundant dictionary to find the non-zeros of the system/signal one at a time until a certain error threshold is satisfied. This approach is named Matching Pursuit [75]. and will be examined in more details in the next section.

3.2 Matching pursuit

In this section, the basic theories of Matching Pursuit (MP) is explained. As explained in the previous section, MP is one possible solution to optimize the process of exploiting the sparsity of the signal by finding the non-zero components iteratively using a redundant matrix and

greedy algorithm. First, we briefly address the greedy algorithm as well as its strength and weaknesses in terms of accuracy and computation load. We then explain the structure of a Matching Pursuit algorithm. Lastly, a modified and improved version of Matching Pursuit, the Orthogonal Matching Pursuit algorithm is introduced.

3.2.1 Greedy algorithm

Greedy algorithm is an algorithm model that based on solving combinatorial problem step by step using a decision making that aims for a local optimum at every step. From the definition, it can be seen that the optimal solution may or may not coincide with the greedy solution and hence the usage of the method should be utilized in very specific scenario and cannot be adaptive to any given problem. This is why sometimes greedy algorithm is attached to the term heuristic instead of algorithm as it provides an approach of learning and discovery of the system rather than right away propose an optimal solution. Nevertheless, Greedy methods, when used correctly, may provide one of the quickest and least expensive way to reach an optimum.

Cases where a greedy paradigm could perform efficiently on usually consists these two properties:

- Greedy choice property: An overall optimum can be achieved by selecting the best choice at each stage. This, in overall, means two things. The first one is that at each stage, the method will only take into account its current stage as well as the cost function or optimal selection in the past to determine its next choice. The second, which can be considered as the most important identity of Greedy, is that regardless of the outcomes or choices that it faces at the next stage, it will never reconsider its past choices. This identity makes it distinctly different from dynamic programming.
- Optimal substructure property: A problem contains optimal substructure property if its optimal solution can be formed from the optimal solution of its sub-problems (problems at every stage).

Depending on the problem, greedy algorithm can be formulated differently. However, in general, these design process always involves these two steps:

- Breaking the given problem into sub-problems that form a sequence of decisions. Depending on the given problem, this step might or might not be required as some problems pose themselves in a sequential way from the beginning.
- Create a general rule for what can be qualified as the optimal choice for the sub-problems. This is the most important feature of the design stage as it directly affects the outcomes of the targeted optimal solution. Sometimes, the rule can be exactly the same as the one for the original problem. However, in other cases, choosing an appropriate



Figure 3.2 – The activity scheduling problem (left) and its proposed Greedy optimal solution (right) which sorts all activity by earliest finishing time and then perform a Greedy choice at each stage based on this criteria. The chosen ones are in blue and the eliminated choices are in red.

qualification rule can be extremely crucial. This highlights the heuristic sense of the Greedy method as there might be multiple way to apply Greedy method to a problem that can result in vastly diverse outcomes.

An example of where the Greedy algorithm thrives is the activity scheduling problem:

- Problem: A set of n activities $A = \{A_1, A_2, A_3, \dots, A_n\}$ with their respective absolute beginning and ending time (t_{b_i}, t_{f_i}) is proposed. Within a bounded time range, choose from them a subset with the highest number of non-overlapping activities. An easier way to imagine this problem is how to fit the most number of classes in a single classroom given their fixed starting and ending time.

Without going into too much details, the correct greedy solution that will lead to a proven global optimum is to pick the activities based on their earliest finishing time. The list of activities is first sorted by earliest finishing time, then at each stage, we always choose the non-overlapping activities with the earliest finishing time out of the remaining available activities. Figure 3.2 summarize the problem and the optimal solution by greedy algorithm.

It can be clearly see from the way we formulate the solution that this is not the only way to come up with a Greedy solution. Other criteria such as shortest interval, earliest starting time, fewest conflicts can be chosen instead but will not lead to a global optimum. This

further highlights the heuristic side of the Greedy method. Nevertheless, the use of Dynamic programming to solve this problem would result in a $O(n^3)$ computational cost. This is why in many scenarios, Greedy algorithm is a preferred choice as it is generally not difficult to design a Greedy method. Furthermore, its run time can always be bounded and it is mostly guaranteed to reach a solution under a reasonable amount of steps.

3.2.2 Matching pursuit

Coming back to the problem regarding an under-determined system in Section 3.1. Other than analyzing the problem at hand by using the ℓ_0 norm equivalence, in [75], a method called Matching Pursuit was proposed where instead of solving the original search for the most sparse solution, it uses a greedy method search process on an over-complete set of dictionaries for the sub-optimal solutions (atoms) that aims to reduce the approximation error.

To formulate the problem under the matching pursuit regime, imagining we are trying to find the most sparse representation of the signal $f \in \mathbb{C}^N$ which is the signal space. An over-complete normalized dictionary \mathcal{D} can be proposed. The dictionary is redundant in the sense that it includes at least N linearly independent atoms that can define the basis of the signal space \mathbb{C}^N . The dictionary $\mathcal{D} = \{g_\gamma\}_{\gamma \in \Gamma}$ in which Γ is the index space of the atoms in \mathcal{D} . The matching pursuit algorithm then attempts to solve:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|f - D\mathbf{x}\|_2^2 \\ \text{s.t.} \quad & \|\mathbf{x}\|_0 \leq M \end{aligned} \tag{3.7}$$

in which M is the final number of atoms chosen to represent f . The algorithm will proceed in a Greedy related manner in which, at each of its iterations denoted m , the atom g_{γ_m} is chosen so that it minimizes the approximation error. One way to satisfy this criteria is to find the normalized atom in \mathcal{D} that will provide the highest inner product between itself and the signal. The contribution of the chosen atom is then subtracted from the signal itself to create the new residue R_m which will be treated as the signal of interest for the next iteration. The

full algorithm can be observed below:

Algorithm 1: Matching Pursuit

Input: Signal f and normalized dictionary \mathcal{D} with column g_γ

Output: Chosen indexes γ_m with $m \in [1, M]$ indicating the atoms in the over-complete dictionary \mathcal{D} .

Initialization

$m=0$;

$R_0 = f$;

while *true* **do**

 Find γ_i to maximize $\langle R_m, g_{\gamma_i} \rangle$;

$\gamma_m = \gamma_i$;

$R_{m+1} = R_m - \langle R_m, g_{\gamma_m} \rangle g_{\gamma_m}$;

$m = m + 1$;

if $|R_m| \leq \text{threshold}$ **then**

break;

end

end

It can be seen that the size of M depends on the stopping criteria. At the end we have:

$$f = \sum_{m=0}^{M-1} \langle R_m, g_{\gamma_m} \rangle g_{\gamma_m} + R_M \quad (3.8)$$

It was proven in [45] that matching pursuit converges exponentially for any function $f \in \mathbb{C}^N$. The Matching pursuit algorithm generally converges more slowly as the span of the signal increases as the closer the coherence vector is to 0 the slower it will approach it according to Jones theorem [59].

One thing that should be noted is that depending on the physics and characteristics of the problem and the concerning signal/function, the stopping criteria could differ. For example, for a unknown system, the minimum requirement of the residual is an appropriate condition. On the other hand, if the signal is known to be K -sparse, a limit could then be imposed on the number of iteration instead to limit the effects of reconstruction noise.

To compare Matching Pursuit with Fourier transform in a time-frequency signal analysis, Fourier transform will fully extract the global frequency aspects of the signal regardless of the characteristics of the signal even if the signal of interest is made of only a small number of distinct sinusoidal. Matching pursuit in this case, can adapt to the signal and only extract the sinusoidal components that best match with the signal. This has an even more significant meaning as Matching pursuit can potentially be used in extracting the physics of the system

that produces such signal.

By using a Greedy foundation, Matching Pursuit can be expected to terminate after a limited number of steps. Once again, due to the greedy nature, the application of matching pursuit should always take into account the optimal structure property. Since its foundation in 1993, Matching pursuit has found its way into multiple applications in 3D object coding [110], image classification [12], video coding with low bit-rate [87, 10], neural networks [92].

Furthermore, various modified version of the original matching pursuit has since been introduced such as Orthogonal Matching Pursuit (OMP), Multi-channel Matching Pursuit [114], Compressive sampling matching pursuit (CoSaMP) [86]...

In the next section, we review a modified version of matching pursuit called Orthogonal Matching Pursuit (OMP) and explain how it can improve MP in terms of reducing bias in the estimation.

3.2.3 Orthogonal Matching Pursuit

In the original Matching Pursuit, the atoms in dictionary \mathcal{D} are generally not mutually orthogonal. Therefore, the update step of the residual, i.e:

$$R_{m+1} = R_m - \langle R_m, g_{\gamma_m} \rangle g_{\gamma_m} \quad (3.9)$$

can possibly introduce components that are not orthogonal to the overall span of the already chosen atoms. This can complicate the convergence process especially in case the level of sparsity for the signal is not entirely known.

In order to eliminate this bias, an orthogonalization projection was studied by Zhang and Mallat to guarantee that the Matching Pursuit algorithm can converge within a bounded number of iterations. This was suggested in [115, 89, 32] that using a Gram-Schmidt algorithm procedure to direct the projection can improve MP significantly.

As mentioned before, if we take a selected atom a non-specific iteration m for instance, g_{γ_m} is generally not orthogonal to g_{γ_i} with $i \in [0, m-1]$. Consequently, when the projection of residual R_m onto g_{γ_m} got subtracted, new components will be introduced onto the space spanned by g_{γ_i} 's. In order to avoid this, the residues can be iteratively projected on an orthogonal basis b_i (with $i \in [0, m-1]$) of the previously chosen atoms.

Without loss of generalization, the first chosen atom is chosen as the first basis b_0 of the atoms' space. At an arbitrary iteration m , the chosen atom g_{γ_m} can be first orthogonalized accordingly

with respect to the space spanned by past chosen atoms using:

$$b_m = g_{\gamma_m} - \sum_{i=0}^{m-1} \frac{\langle g_{\gamma_i}, b_i \rangle}{||b_i||^2} b_i. \quad (3.10)$$

Due to this orthogonalization of the newly selected atom, the projection stage for the residue is also going to be changed into a projection onto b_m instead:

$$R_{m+1} = R_m - \frac{\langle R_m, b_m \rangle}{||b_m||^2} b_m \quad (3.11)$$

Assuming that the algorithm terminates after M iterations, the formulation of original signal f becomes:

$$f = \sum_{m=0}^{M-1} \frac{\langle R_m, b_m \rangle}{||b_m||^2} b_m + R_M \quad (3.12)$$

which is basically the projection of f onto the space spanned by all the chosen atoms plus the remaining residue. All of the convergence conditions from the basic MP holds for OMP. Needless to say, at the end, f needs to be expressed on the basis of the original g_{γ_m} 's basis. This can easily be done by inverting all the b_m back using a transformation matrix based on

Eq. 3.10. The summary of the method is presented below:

Algorithm 2: Orthogonal Matching Pursuit

Input: Signal f and normalized dictionary \mathcal{D} with column g_γ

Output: Chosen indexes γ_m with $m \in [1, M]$ indicating the atoms in the over-complete dictionary \mathcal{D} .

Initialization

$m=0$;

$R_0 = f$;

while *true* **do**

 Find γ_i to maximize $\langle R_m, g_{\gamma_i} \rangle$;

$\gamma_m = \gamma_i$;

 Orthogonalize g_γ according to the space spanned by past chosen atoms

$$b_m = g_{\gamma_m} - \sum_{i=0}^{m-1} \frac{\langle g_{\gamma_i}, b_i \rangle}{\|b_i\|^2} b_i ;$$

 New residue is formed by subtracting current residue by its projection on the orthogonalized atom b_m ;

$$R_{m+1} = R_m - \frac{\langle R_m, b_m \rangle}{\|b_m\|^2} b_m ;$$

$m = m + 1$;

if $|R_m| \leq \text{threshold}$ **then**

 break;

end

end

If needed, inverting back to original atoms coordinates using a transformation of basis.

It was shown in [45] that MP and OMP tends to perform in an identical manner for the first few loops, this is due to the fact that MP also tends to extract atoms that are mutually orthogonal for the first few iterations. When the number of iteration increases, OMP will typically have a smaller residue than compared to MP. The orthogonal projection and back projection of OMP significantly increases the cost of OMP compared to MP. However, OMP has a higher bound for accuracy than the basic MP. Another general comments on both MP and OMP is that for large scale problems, the correlation calculation can also become costly. In order to reduce the processing time, technique using the discrete Fourier transform as in [49] can also be used effectively if the atoms in the dictionary is equi-spaced frequency wise.

OMP has been widely used since its emergence in a diverse field of applications. This includes high dimensional sparse signal recovery [19], recursive function approximation as well as signal recoveries from random measurements [111]. There have also been modified version of OMP to suit different domains and objectives such as Stagewise OMP (StOMP) [35], Gen-

eralized OMP (gOMP) [113] as well as Simultaneous orthogonal matching pursuit (SOMP) [25].

3.3 Conclusion

In this chapter, we have briefly reviewed the theme of Sparse representation and Sparse recovery. The objective is to understand how inherently sparse system or signals could be reconstructed/recovered using a much smaller number of measurements/sampling points. Through the analysis of the inherent sparsity of the system or signal, it has been shown that the entire sparse recovery problem centers upon the very concept of a ℓ_0 norm optimization. This problem, in its entirety, is proven to be highly difficult and expensive to solve using the direct route. This is why modern research has been coming up with novel methods to tackle the original problem. These methods, more often than not, take one of the two possible routes. The first one is based on the argument that solving an ℓ_0 norm minimization problem is very much similar to solving the equivalent ℓ_1 norm problem. Several algorithm has been developed based on this idea such as basis pursuit and compressed sensing. The other route tries to take a very much different point of view on the problem through the scope of a matching process through Greedy algorithm. By making use of a overly complete dictionary, the method, at every iteration, attempts to search for the best atom in its dictionary that best match with the system or signal through the use of the correlation function. By progressing iteratively on a Greedy basis, Matching pursuit can ultimately choose the best atoms in its dictionary that matches best with the signal. An improved version of Matching pursuit which incorporated a Gram-Schmidt orthogonalization process into the original algorithm is called Orthogonal Matching Pursuit or OMP. By using this orthogonalization of the entire matching system at every iteration, the results could converge quicker when compared to the original one.

Looking back at chapter 2, we have analyzed how the sound field in room, at its core, is composed of the sum of its discrete modal responses. This means that if we have an over-complete dictionary that is somehow spread on the frequency axis, only a few of these atoms would be highly matched with responses in the room. These atoms will likely be linked to the eigenmodes of the room. This is a highly convenient condition for the use of a Matching Pursuit-based algorithm. Moreover, taking some assumptions into consideration, the eigenfunctions, or equivalently, the mode shape functions of the room are also mutually orthogonal (see Eq. 2.5) which benefit both the Greedy algorithm and the orthogonalization process. Using these reasoning, the next chapter will be devoted entirely to analyzing the spatial response of the room and how they can be reconstructed using a small number of measurement points with the help of a modified version of the Orthogonal Matching Pursuit.

4 Sound field reconstruction framework

One of the main motivation of this research is to propose a practical, yet precise reconstruction framework for the sound field in a room at low frequencies. As can be seen from Section 2.2, existing measurement methods in the room are mostly concerning its time or frequency representation while the spatial information has been proved difficult to recover. A thorough recovery of the enclosed sound field has been shown to use a large number of microphones, often goes up to around a hundred measurement points to maintain a stable level of accuracy.

In this chapter^I, with the help of the sparse recovery schemes which have been briefly introduced in Chapter 3, we aim to create a practical framework for the recovery of sound field for a large area inside a non-rectangular room. The practical aspect of the framework involves reducing both the number of required measurement points (which is directly related to the number of microphones) and the computational time (which is equivalent to the complexities of the system) without sacrificing the accuracy of the framework. As the accuracy of the reconstruction is also an essential part of the framework, irrespective of the methods that could be used to reduce the number of measurement points, such methods should substantially rely on a sparse representation of the wave equation. In this chapter, we address the wave solution for the sound field in a room at low frequencies which has been discussed in Chapter 2. Through different decomposition and approximation methods, several sparse properties of room acoustics can be brought to light in hopes of reducing the complexities of the system. These properties can be categorized into either exact sparsity, which are features that inherently emerge from the physics of the room, or approximate sparsity, which requires an approximation process to reduce the number of degrees of freedom in the wave equation. The ultimate objective is to obtain a governing equation of the spatial response of sound pressure in the room where the number of variables can be well defined and quantifiable. After obtaining the closed form of the governing equation, the structure of the framework will be discussed in details with the aim to recover the unknown parameters from a small number of measurement points. The reconstruction framework detailed in this chapter will serve as

^IChapter 4 and 5 are partially based on the excerpts of "T. Pham Vu et al., *Acta Acustica* 4(2), 2020; <https://doi.org/10.1051/aacus/2020006>"

the foundation for the reconstruction procedures both for simulation and experimental data that follows in Chapter 5.

4.1 Sparsity in room acoustics at low frequencies

Depending on the region of the wavelengths of interest, room acoustics can be analyzed and studied using different approaches. At low frequencies, where the wavelengths are of the same order as the dimensions of the room, the acoustic behavior of the sound in the room can be explained through the reflection of sound waves by the reacting boundaries (walls) of the room. This gives rise to the standing waves phenomenon in the room which are produced from the interference between the incident and reflected waves at certain frequencies called eigenfrequencies. At low frequencies, these room modes are where most of the acoustic energy are concentrated at. Consequently, the acoustic behavior of the room at this frequency range is dominated by the responses of these discrete room modes [65].

4.1.1 Modal decomposition

At low frequencies, there exists a formulation of the acoustic response in the room, where, at any frequencies, the wave solution can be decomposed as a discrete sum of damped harmonic responses of eigenmodes.

$$p(t, X) = \sum_n A_n \Phi_n(X) g_n(t) \quad (4.1)$$

where X is the location of the point of interest, $\Phi_n(X)$ are the space-dependent mode shape functions (i.e, the eigenfunctions of the Helmholtz equation) for each mode n of the room at location X , $g_n(t)$ is the harmonic time-dependent decaying function for each mode and A_n are the corresponding complex expansion coefficients for each of the mode. This form of decomposition can be acknowledged as a sparse representation of the room acoustic response since it transforms the response at any point in the room to a discrete sum of damped harmonic eigenmodes. Figure 4.1 demonstrates the modal decomposition/summation characteristics through an example of a Room Frequency Response (RFR).

In this representation, each mode of the room is uniquely presented by a complex wavenumber $k_n = (\omega_n + j\delta_n)/c_0$ which can also be referred to as the eigenvalues of the Helmholtz equation, where c_0 is the speed of sound in the air, ω_n is the modal angular frequency and $\delta_n > 0$ is the corresponding modal damping [84]. The harmonic time dependent decaying function $g_n(t)$

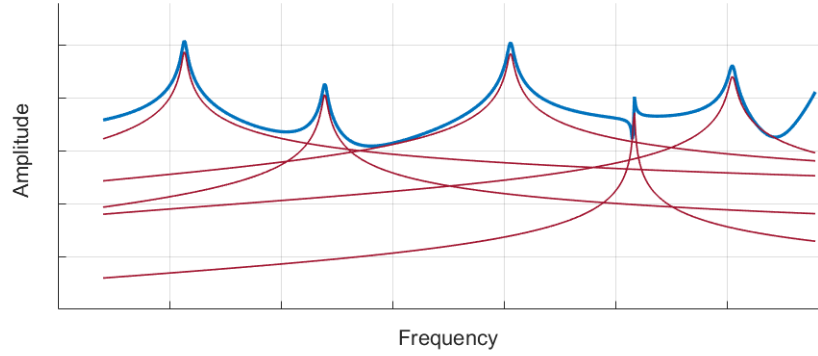


Figure 4.1 – Modal decomposition/summation for an RFR curve. The response can be seen to be composed of discrete modal responses.

can then be expanded as:

$$g_n(t) = e^{jk_n c_0 t} = e^{j(\omega_n + j\delta_n)t} = e^{j\omega_n t} e^{-\delta_n t} \quad (4.2)$$

The mode shape function Φ_n , on the other hand, is a space dependent function. For each mode n , the mode shape function details the spatial pattern of the acoustics in the room when it is excited at the corresponding eigenfrequency. In case of ideally rigid walls, each mode shape function is the exact eigenfunction of the Helmholtz equation [48]:

$$\Delta\Phi_n + k_n^2\Phi_n = 0 \quad (4.3)$$

where k_n is the eigenvalue of mode n . In the case of non-rigid walls, it can be assumed that the above equation still remains valid for cases where the domains of interest are not too close to the wall as well as for the case where the walls are not highly damped.

It is worth highlighting that X is referring to the spatial location of the receiver/microphone of interest while the location of the source (or sources) in the room are not present in equation 4.1. In fact, the information regarding the source, though not visibly seen, has been accounted for within the complex coefficients A_n and will remain implicit all following derivations. This is prompted by the fact that under the scope of our investigation, only a single fixed source will be considered and hence the terms corresponding for it remain constant and can be considered as a part of the expansion coefficients.

4.1.2 Mode shape approximation

From the previous expansions, we have been able to observe the inherent structured sparsity of the wave solution through the modal decomposition form where the response at a location in a room is decomposed into a discrete sum of modal response of room modes. From equation 4.1, it can be seen that although the positional vector \vec{X} is present, this is not yet a closed form solution from the reconstruction problem point of view. In order to have a complete reconstruction/recovery form of the acoustics in a room at low frequencies, there requires another process to obtain a universal expansion for the mode shape functions of the room. In this section, we address this issue through the use of the mode shapes approximation techniques.

There have been various attempts to achieve an expansion form of eigenfunctions in general [18, 47, 63], and of the mode shape functions in particular. In [82], it has been shown that using the premise of spherical harmonics and spherical Bessel functions, the mode shape functions can be generalized using an approximation theorem. According to this, the mode shape function can be approximated by a finite sum of plane waves sharing the same wave number but pointing in different directions. To be more specific, each individual mode shape, hence, can be formulated under the following R -th order approximation:

$$\Phi_n(\vec{X}) \approx \sum_{r=1}^R B_{n,r} e^{j\vec{k}_{n,r} \cdot \vec{X}} \quad (4.4)$$

of which $\vec{k}_{n,r}$ can be expressed as the 3D wavevectors sharing the same wavenumber $\|\vec{k}_{n,r}\|_2 = |k_n|$ (i.e, sharing the same vector length) and $B_{n,r}$ is the expansion coefficient for each wavevector. One thing that needs to be clarified is that in opposition to the inherent sparsity in the previous section, this is an approximate sparsity. This decomposition not only provides an approximation for each of the mode shape, but on a wider scope, it also exhibits a closed-form interpretation of the mode shape function regardless of the type of the modes in the room. Assuming that we now restrict this representation below a certain upper frequency limit, a finite number R of wavevectors would be sufficient to closely approximate each and every mode shape function within this frequency range. Combining equation 4.1, 4.2 and 4.4, we achieve the following expansion:

$$p(t, \vec{X}) = \sum_{n,r} C_{n,r} e^{j\omega_n t} e^{-\delta_n t} e^{j\vec{k}_{n,r} \cdot \vec{X}} \quad (4.5)$$

where $C_{n,r} = A_n B_{n,r}$ with $r \leq R$. Hence, through a series of expansion, the expression in Eq. (4.1) can be interpreted as the discrete sum of space-time damped harmonics with the expansion coefficients $C_{n,r}$. This decomposition form directly links the acoustic response of

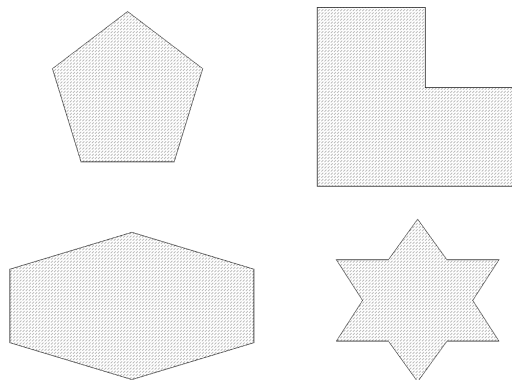


Figure 4.2 – Possible star-shaped polygons that satisfy the condition for the recovery of mode shape functions

the receiver to its location in a room.

4.2 Reconstruction framework

Once we have arrived at the close form expression of the spatial response in the room, it becomes clear that the objective of the reconstruction framework is to identify and estimate the values of the unknown parameters in Eq. (4.5) from a limited set of measurements. First of all, we should clearly identify the known and unknown parameters. In this research, one of the goal is to approach this reconstruction problem in the most generalized way, i.e assuming that very little number of information would be available at hand. The mode shape approximation technique in Section 4.1.2 only requires the room to be a star-shaped polygon. This means that there exist at least a point in the room that can reach out to every other possible point in the room. This geometry requirement in reality is quite tolerant in a sense that most existing rooms would satisfy except for a few cases of connecting rooms in a floor plan. Fig 4.2 listed a few example of 2D polygons that could easily satisfy this requirement. From this motivation, in this research, we assess the general case of a non-rectangular room where the modal behaviors of which are much less predictable than in a traditional shoe-box room. Figure 4.3 shows the geometry of a non-rectangular reverberation chamber with two simulated room modes. Notice that due to the irregular architecture of 3D angled walls, even though the room has a conventional number of walls, the spatial behavior of the mode shapes in the room are extremely difficult to predict, even for familiar types of mode such as the axial mode [62, 108].

Coming back to the reconstruction problem at hand, the inputs of the framework are clearly the microphone measurements and their 3D location in the room. Inside the room, assuming a number of M microphones are randomly placed at different locations to acquire the room responses. As we are dealing with the low frequency range, these measurements could be

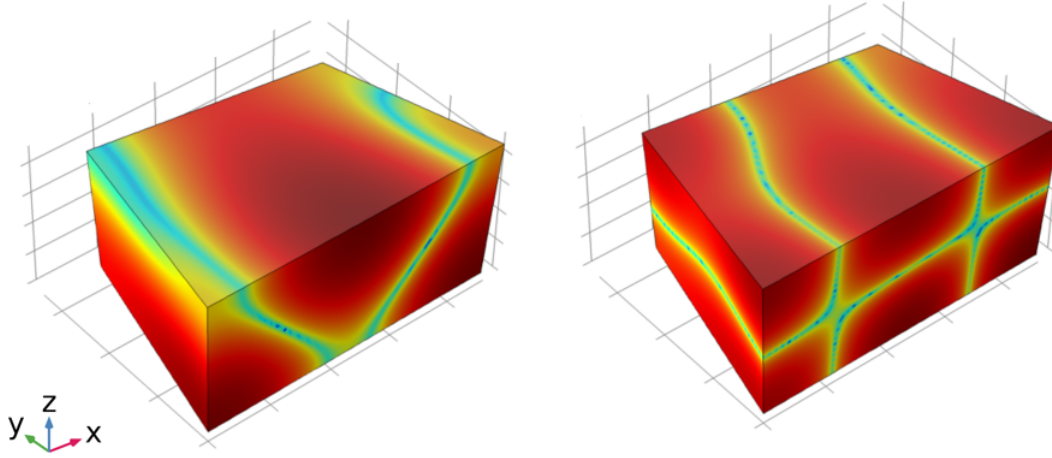


Figure 4.3 – Examples of room modes in a non-rectangular room.

filtered as well as downsampled to reduce computational cost. Calling N_t the length of the time vector of each microphone measurement, the $(N_t \times M)$ matrix \mathbf{S} of signals is defined as the acoustic measurement matrix. This, in combination with the M positional vectors \mathbf{X}_m of the measurement points, establish the inputs for the reconstruction framework.

The outputs of the framework, in essence, should be all the unknowns present in Eq. (4.5), excluding the predefined parameters, namely, the number of modes N and the list of wavevectors $\vec{k}_{n,r}$ (which creation will be discussed later) for each mode shape approximation. Considering the fact that we assume to not know much about both the geometry and the absorption characteristic of the walls, the outputs, hence, include the angular frequency ω_n and the exponential damping factor δ_n for each eigenmode. Along with that are the $N \times R$ expansion coefficient $C_{n,r}$ for each of the plane waves term. Once all of these unknown parameters are found, it is possible to interpolate the responses at any position \mathbf{X}_{int} in the room simply by plugging it into Eq. (4.5).

The detailed framework can be divided into two steps. The first step, detailed in Section 4.2.1 is called modal identification, aiming at estimating the modal wavenumbers k_n for the N room modes. This includes identifying both the frequency and damping for every concerned modes of the room. Once these parameters are identified, the second step, detailed in Section 4.2.2 focuses towards approximating the expansion coefficients $C_{n,r}$ for a set of predefined wavevectors $\vec{k}_{n,r}$ through projection.

4.2.1 Modal identification

In this section, two different approaches are addressed, processing the input signals either in the time domain or in the frequency domain. The first approach is the Simultaneous Orthogonal Matching Pursuit method [111, 25] (SOMP) which is based on the original Orthogonal

Matching Pursuit (OMP) that has been introduced in Section 3.2.3. This method is a modified version based on a greedy algorithm approach to recursively estimate modal parameters for each eigenmode of the room from the matrix of input time signals. The second method is based on the Rational Fraction Polynomials (RFP) global curve fitting method [52] which, in contrary to the iterative SOMP, simultaneously estimates the modal parameters of the room from a set of input Room Frequency Responses of the room. Each of the two method has its own advantages and disadvantages depending on various factors such as accuracy, robustness as well as processing time. Both methods will be assessed and used in the reconstruction procedure that comes after.

Time domain approach

In this section, we analyzed a modified version of the Orthogonal matching pursuit algorithm [111, 25, 79] where the residue is a matrix of all the measurement signals. This method aims to performed an iterative modal identification on the temporal matrix of measurement system. In short, from a redundant set of pre-defined damped sinusoids, this method can perform a search for the ones that are highly correlated with the input signals measurement matrix using a Low-rank Greedy approximation approach.

At the beginning, two different sets of ω and δ with $\omega_{min} < \omega < \omega_{max}$ and $\delta_{min} < \delta < \delta_{max}$, are constructed. The variation range for these sets are roughly estimated based on available knowledge on the room including the general dimensions of the room as well as some test run of modal decay times. The combinations of every pair of entries from the two sets together will then assemble an over complete set of complex components ($j\omega_q - \delta_q$) in which $q \in [1, Q]$ with Q as the total number of possible combinations. This section of the framework represents a very common problem design of a Matching Pursuit algorithm. The significant difference here from a regular Fourier dictionary is that the dictionary contains not just regular sinusoids but damped sinusoids with a complex angular frequency. Each combination from this set with the time vector can then be used to produce a time vector with length N_t of time-decaying damped sinusoid:

$$\theta_q = e^{j\omega_q t} e^{-\delta_q t} \quad (4.6)$$

Using the normalized vectors of $\overline{\theta}_q = \theta_q / \|\theta_q\|_2$ as column vectors will give provide a $(N_t \times Q)$ array $\overline{\Theta}$. This array can be considered as a normalized dictionary of time decaying damped sinusoids. This normalization or non-dimensionalization is also a commonly encountered step of both MP and OMP (refer to Section 3.2.2).

The matching algorithm, in short, will use this dictionary to repetitively perform a pole searching procedure on the measurement matrix in iterations. Similar to the traditional MP, every loop indexed i starts with assigning a $(N_t \times M)$ residue matrix \mathbf{R}_i which is the result of

the previous loop. At the first loop, \mathbf{R}_1 is set to be equal to the predefined measurement matrix \mathbf{S} . Through the searching and matching procedure, a damped sinusoid from the dictionary with the highest correlation to the residue matrix is chosen. This selected damped sinusoid is representing a pair of ω_n and δ_n which are the modal parameters of one of the room modes. The new residue matrix \mathbf{R}_{i+1} for the following loop can then be formed by extracting the contribution of this chosen damped sinusoid from \mathbf{R}_i . The algorithm for a non-specific i^{th} iteration can be detailed below:

- Construct the $(Q \times M)$ correlation matrix $\Xi_i = |\bar{\Theta}^H \mathbf{R}_i|$. This correlation matrix contains all the correlation information between the dictionary and the residue matrix (which at the first loop is the measurement matrix itself). Each row q of Ξ_i contains a number of M correlation values between the q^{th} normalized damped sinusoid from the dictionary and each of the M columns of the residue matrix.
- By summing the energy of this set of M values, compute the evaluation correlation value σ_q between the q^{th} damped sinusoid and the entire set of measurements: $\sigma_q = \sum_{m=1}^M (\Xi_{i[q,m]})^2$.
- Out of the Q available values of σ_q , select the maximum one, which corresponds to the pole with the highest correlation to the measurements/residue.
- As a result, the identified index (let's call it q_i) will provide the chosen modal wavenumber of this loop, which is: $k_i = (\omega_{q_i} + j\delta_{q_i}) / c_0$.
- After the modal parameters for this loop is found, following the orthogonalization and projection of SOMP in Ref. [111], the residue matrix to start the next loop can be computed by eliminating the contribution of the chosen sinusoids of this loop from the current residue: $\mathbf{R}_{i+1} = \mathbf{R}_i - \mathcal{P}_i \mathbf{R}_i$ in which \mathcal{P}_i is the orthogonal projection onto the previously chosen damped sinusoidal.
- Repeat with $i = i + 1$ until $i = N$

At the end of the iterative processes, a group of N complex wavenumbers corresponding to the eigenmodes of the room is established. As mentioned in Section 3.2.3, the orthogonalization process, though can seem expensive, does result in a faster convergence of the algorithm. Furthermore, due to the mutually orthogonal modes, the orthogonalization prevent adding the same damped sinusoid from the dictionary more than once. Another process that on the surface seems expensive is the summing of correlation process. This can be improved by defining an upper bounded equally spaced set of ω for the dictionary and make use of a process similar to Fourier transform detailed in [49] on the correlation matrix to reduce the dimension of the ranking process.

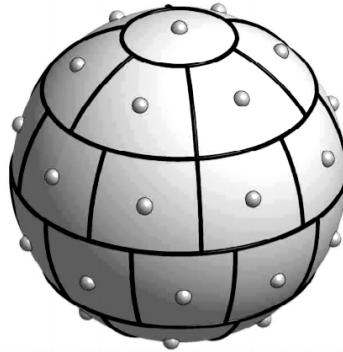


Figure 4.4 – Example of the spherical sampling that can be used to generate wavevectors for the mode shape approximation process by Paul Leopardi [69].

Frequency domain approach

As room modes are usually readily visible in the Room Frequency Responses (RFRs), it is logical to investigate a frequency-domain approach for room modes identification besides the temporal approach. There are several approaches that have been under studied to analyze the modal properties based on the frequency responses [91, 50, 66]. One particular example is the global curve-fitting method in the frequency domain using the Rational Fraction Polynomial (RFP) formulation [52]. This has been used in Ref. [102] to estimate the modal parameters by curve fitting the RFR measurements.

Conventional curve-fitting methods are usually processed locally, initiating the fitting process on a single function at a time. The concerning method in Ref. [52], however, performs curve-fitting procedures on multiple frequency response functions at different locations in a room simultaneously to identify the governing model of the system. The method assumes the linearity of the RFRs and that they can be formulated as a ratio of two polynomials. These RFRs share the same denominator whose poles contain information on the modal angular frequencies (ω_n) and damping (δ_n) of the room. The method then performs a concurrent curve-fitting on the set of measured RFRs to acquire the modal parameters of the room within a given bandwidth. This can be explained in more details in Appendix A.

4.2.2 Projection onto spherical sampled wavevectors

Up to now, only the eigenfrequency parameters of the room modes, namely, ω_n and δ_n given in equation (4.5) have been identified. The remaining parameters to be determined are the expansion coefficients $C_{n,r}$, for which the following algorithm is used:

- The first step is the separation of the current known and unknown parameters. Note that the time-varying terms in Eq. (4.2) have been identified in the former algorithm, and can

be discarded from now on. Using a matrix form accordingly of the measurement matrix \mathbf{S} gives: $\mathbf{S}^T = \Psi G$ with G as the $(N \times N_t)$ matrix where each of its row is a modal damped sinusoidal $g_n(t) = e^{j\omega_n t} e^{-\delta_n t} = e^{jk_n t}$. Furthermore, Ψ is the $(M \times N)$ space-dependent matrix of modes with the inclusion of the expansion coefficients A_n that appear in Eq.(4.2):

$$\Psi_{[m,n]} = A_n \Phi_n(\mathbf{X}_m) \quad (4.7)$$

with \mathbf{X}_m 's the M position vectors for the location of the input measurements of \mathbf{S} . If $N_t > N$ (which usually is the case), the system of Eq.(4.7) is over-determined with $(M \times N)$ unknown and $(M \times N_t)$ equations. Hence, it is possible to estimate the $(M \times N)$ matrix of Ψ by computing the least-squares estimation:

$$\Psi \approx \mathbf{S}^T G^H (G G^H)^{-1} \quad (4.8)$$

- Based on the expression in Eq.(4.5), Ψ can be further expanded using plane waves expansion:
 - * First, the list of component wavevectors needs to be defined. For each mode shape function, a set of R wavevectors $\vec{k}_{n,r}$ is created whose norm and directions match a uniform sampling over a sphere with radius $|\omega_n/c_0|$. Spherical sampling (proposed in Ref. [69]) is chosen in this case because the room is non-rectangular and hence there is no preferred basis for the formation of mode shape functions. An example of this can be seen in Fig.4.4.
 - * Each column ψ_n of the matrix Ψ can be treated individually as they are associated with different modes. Calling ρ_n the $(M \times R)$ matrix of the plane wave harmonics for mode n in which $\rho_{n[m,r]} = e^{j\vec{k}_{n,r} \cdot \vec{X}_m}$, each column vector ψ_n can be individually characterized as:

$$\psi_n = \rho_n \mathbf{C}_n \quad (4.9)$$

with \mathbf{C}_n the $(R \times 1)$ vector consisting of the R expansion coefficients $C_{n,r}$ of mode n . First, assuming that $R < M$, taking ρ_n as the basis, ψ_n can be projected onto this basis to derive the coefficient vector \mathbf{C}_n using least-square projection:

$$\mathbf{C}_n \approx (\rho_n^H \rho_n)^{-1} \rho_n^H \psi_n \quad (4.10)$$

As mentioned above, this derivation is only available when the number of sampled plane waves is lower than the number of microphones. As can be seen in Ref.

[82],[7],[56], the convergence of the plane wave approximation is highly dependent on the number of plane waves available, especially in 3D. Hence, in the case where the number of measurement points is fairly low, restricting $R < M$ could affect the reconstruction of mode shape functions. One possibility would be to allow $R > M$ and derive the coefficient vector using a least norm optimization:

$$\mathbf{C}_n \approx \rho_n^H (\rho_n \rho_n^H)^{-1} \psi_n \quad (4.11)$$

Further studies need to be done to verify the limitations of this solution as well as the optimal choice for R . In our case, for a low number of microphones, several trials have shown that choosing $R > M$ can estimate the mode shape better and increase the overall correlation. Regardless of the method used, in practice, the applicability of this step can always be cross-checked using a number of evaluation microphones.

Repeating the technique on each mode $n \leq N$ will return the set of expansion coefficients $C_{n,r}$ required for the reconstruction.

Through the derivations in Section 4.2.1 and 4.2.2, we have been able to attain the necessary parameters for the reconstruction process. Chapter 5 is dedicated to the assessment of the reconstruction results for a non rectangular room.

4.3 Conclusion

In this chapter, we have analyzed the spatial responses in the room and its governing equation through the scope of Sparse representation that has been suggested from the previous chapter. Through multiple derivations and approximation, we have exploited both the inherent structured sparsity from the spatial response of the room due to the existence of eigenmodes and the approximation sparsity using the mode shape approximation method. With the help of these processes, a closed form equation of the spatial responses in the room can be achieved and is ready to be addressed using sparse recovery. In order to fully recover the sound field within the room, two different approaches were suggested to first recover the modal parameters of the room. These methods includes a modified version of OMP, where a series of signals can be grouped together and form the residue matrix and an entirely different method in the frequency domain that makes use of the Rational fraction polynomials formulation to search for the poles and zeros on a system of responses. After the modal parameters of the room has been recovered, through the use of the spherical sampling of plane waves and projection regime, the expansion coefficients of the terms in the reconstruction formula can be recovered. This final form of the spatial response in the room will potentially allow us to recover the responses in the room at any given points. The next chapter will focus on implementing this framework on both simulation data and real experiments to confirm its accuracy and validity.

5 Reconstruction results

The reconstruction framework in the previous chapter has laid the foundation for the spatial reconstruction of the sound field in a room. Through a series of derivation, a closed form governing equation has been found, linking between the location of a point in the room with its corresponding acoustic response.

In this chapter, we propose to put this framework to test by adapting it for both simulation and real measurements in a room. A simulation is first introduced with the model of a non-rectangular room as the framework in Chapter 4 was not built specifically for just a traditional shoe-box room. This non-rectangular design represents real life situation where not all rooms can be considered as a rectangular room, this is also true for rectangular room with large objects or furniture. The numerical study is done through a Finite Element Method (FEM) which has been introduced in Chapter 2. Through the numerical studies, the reconstruction results can be evaluated first, by assessing the interpolated individual response and furthermore, by a large scale spatial comparison, where the acoustic response of the room can be reconstructed for a large rectangular sub-space. From an initial study of a room with very low wall damping, we can progress upon the sound field reconstruction for damped rooms and analyze the effects it has on the accuracy of the results. In addition, we also tackle the analysis of different aspects regarding the number of input measurement points as it is directly linked to both the robustness and the practicality of the framework. These in-depth analysis on the simulation studies will provide us with valuable information before moving on to the experimental studies which are detailed in Section 5.2, where real measurements are performed in a reverberation chamber with the same design as the model used for FEM. The reconstruction results can then be evaluated to confirm the validation of the framework.

5.1 Numerical simulation

In this section, we start with the analysis of the reconstruction framework using a numerical approach. This first numerical study allows the assessment of multiple aspects of the reconstruction framework. The numerical approach provides access to a very fine distribution of

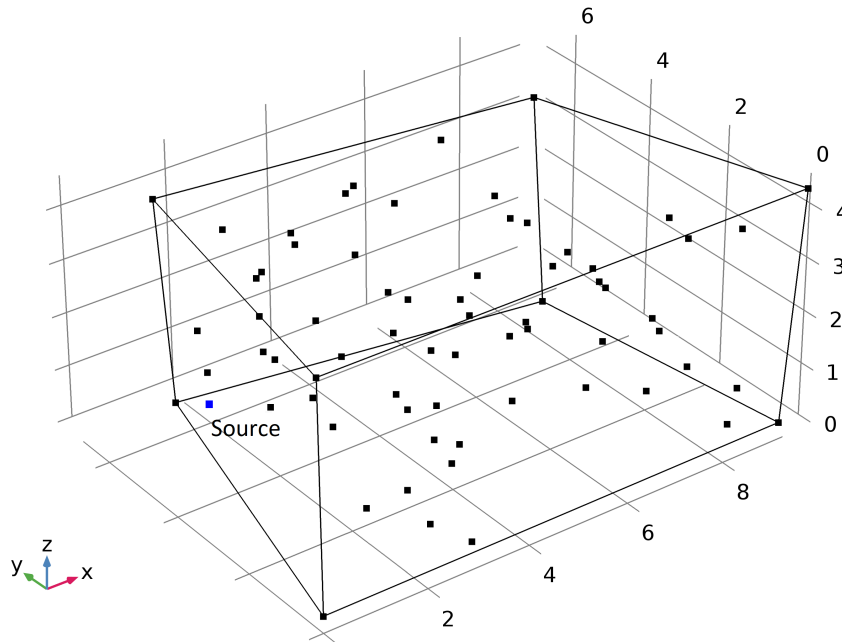


Figure 5.1 – Geometry of the FEM model. The black dots represent the measurement points that are spread randomly in the room.

microphone placements which allows a broad range of examination on numerous aspects of the reconstruction results. Furthermore, the properties of the room, most significantly, the wall absorption, can be modified straightforwardly. The most crucial advantage of the numerical study is that the FEM simulation not only provides the input but also can be used as the ground-truth reference for cross-checking the reconstruction results on a universal scale.

We begin with the FEM model of a non-rectangular room using the COMSOL simulation software. The FEM model is consisted of a non-rectangular room with a maximum height of 4.6m, maximum width of 9.8m and maximum length of 6.6m that replicates the actual reverberation chamber exists at École Polytechnique Fédérale de Lausanne (EPFL). The damping of the walls are initially set to be very low, with a uniform absorption coefficient of $\alpha = 0.01$ which results in a real valued reflection coefficient, approaching that of the actual reverberation chamber. The sound source is simply designed as a monopole point source and is placed very closed to a corner of the room. This gives the source the ability to effortlessly excite all the modes of the room. The measurement points, on the other hand, are spread randomly in the room (refer to Figure 5.1). This particular placement method, although perhaps not the most ideal microphone placement strategy for a specifically given geometry, is however the most guaranteed to capture sufficient information regarding the sound field and its modal properties, especially in the general cases where it is assumed that there are no readily available knowledge on the room properties. Figure 5.1 showcase the FEM model of the room along with the randomly generated measurement points.

5.1.1 Modal identification

In this section, we first performed modal identification based on the measurements from the multiple receivers. As mentioned in Section 4.2.1, two different methods are under studies, namely SOMP and RFP methods. Modal identification is a crucial process which can directly affects the accuracy of the reconstruction in the recovery stage. The objective of the identification process is to search for the modal parameters of the room, including the modal frequency and modal damping. In this analysis, however, instead of directly comparing the retrieved modal properties, namely ω_n and δ_n , the focus has been put on two other equivalent but more practical properties in modal analysis: the eigenfrequency (f_n) and the modal decay time (MT_{60n}) - which can be defined as [105, 9]:

$$MT_{60n} = \frac{3 \ln(10)}{\delta_n} \quad (5.1)$$

These two practical properties reflect the modal properties of the room and are directly linked to ω_n and δ_n . Using the same number of 40 microphones, the modal decay times estimated from SOMP and RFP methods for the first 12 modes of the room are compared with those computed from the baseline FEM analysis which can be considered as the ground-truth (see Figure 5.2). Note that, as of this moment, not much consideration has been given into the number of microphones because the focus is solely on how both methods perform considering they received the same number of inputs. As can be proven later in the next section, the number of required microphone can be lower than this.

Through a few initial trial tests, it was acknowledged that both methods generally performed equally well in identifying the eigenfrequency f_n (in Hz) for each of the modes of the room. As the values of f_n obtained using the two methods do not present much discrepancy, the comparison in Fig. 5.2 is demonstrated here only in terms of the modal decay times to examine their performance with respect to modal damping estimation. Using the numerical results from the FEM analysis as the reference, it can be seen in Figure 5.2 that both the RFP and SOMP methods are capable of identifying the room modes damping, except that SOMP, on average, may underestimate the damping for the mode at 40.5 Hz, which will be discussed later on in the section.

Overall, it can be observed that the RFP identification method performs marginally better than SOMP. However, it is worth noticing that there are significant differences between the two methods regarding the robust aspect. Even though both methods need a manual prerequisite input concerning the total number of modes within a limited bandwidth, they each process this information differently. For the global curve fitting method using RFP, if the total number of modes within the defined frequency limits is not accurately known beforehand, usually a considerable amount of trials and errors are necessary to eventually come up with a coherent curve-fitting result. Moreover, it can be seen at the later stages that without a meticulous check

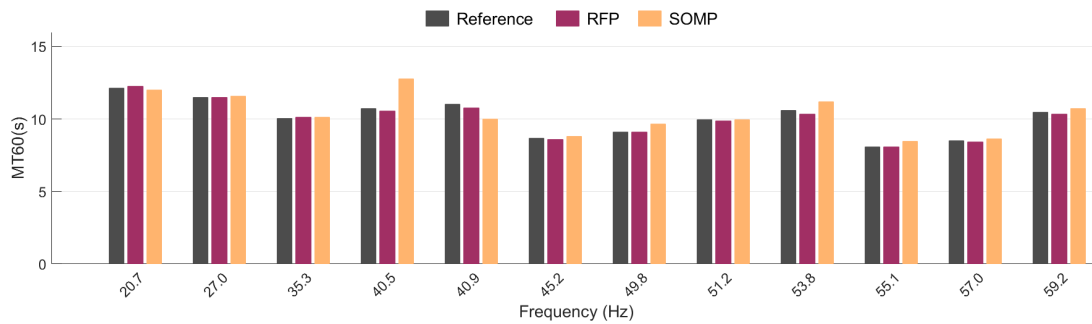


Figure 5.2 – Modal decay times for the eigenfrequencies of the non-rectangular reverberant room as estimated by the SOMP and RFP methods, in comparison with the FEM analysis (reference).

of consistency, the interpolation results using RFP could end up with a higher amount of errors. This vulnerability and instability, for most cases, does not exist for SOMP. The main reason behind this robustness is that the modal parameters are found in RFP simultaneously whereas in SOMP they are found individually through the repetitive loops. Furthermore, for SOMP, this iterative procedure progresses in a residual manner: the eigenmodes that have the highest contributions in terms of energy to the collected signals are located first, followed by the ones with less contribution. This gives SOMP an advantage especially during the reconstruction process as the results do not deviate much from reality when the number of modes within the frequency of interest are under/overestimated. The actual number that users enter into SOMP can only change the number of times the algorithm is repeated but should not affect the result of each individual loop.

In the case of Fig. 5.2, the underestimated damping by SOMP that sometimes occurs at 40.5 Hz also comes from the fact that the concerning algorithm processes the residue at the end of each computing step. At the end of each iteration, the contribution of the found eigenmode is taken away from the present residue by the help of a projection matrix to create a new one for the next loop. Consequently, the modes that are found at the later iterations of the algorithm are prone to higher errors and also, its correlation with the measurements is likely to be less than the ones that come before in the algorithm. When there are two modes that are very close together such as the particular cases at 40.5 Hz and 40.9 Hz respectively, depending on the set of input measurements, one of them may be found at the very far end of the algorithm compared to the other. Since one of these modes has been identified earlier in the process, and its contribution to the residual has been extracted before, the error that occurs at the other would have minimal effects to the overall reconstruction result in the next stage. The same situation also occurs when users overestimate the number of modes. A redundant number of modes means that at around the final loops, the algorithm will inevitably find some frequencies that do not correspond to any of the room modes. As long as the overestimation is not too far from reality, this error in SOMP would only have minor effects on the reconstruction

results because the contributions of the few wrongly-found modes are generally insignificantly small when compared to the correct ones.

Albeit being less robust, the RFP curve fitting method does have a clear advantage over the SOMP method regarding the overall computational cost. Not only that SOMP performs a repetitive mode finding process that regularly requires the update of the residual matrix but also performs it using multiple costly matrix operations. The RFP method developed in Ref. [98] on the other hand, does not perform an iterative process and has taken into account several computational simplifications. For instance, on a conventional work station with 32GB RAM and 4 cores CPU of 3.4 GHz, in order to accomplish the computation that return the results in Figure 5.2 using 40 microphones, the SOMP method would usually take 4-5 minutes to finish while the RFP method would finish in 10 seconds. This significant difference will further increase if the number of input measurements or the number of modes increases.

Overall, it can be concluded that SOMP is a robust method that works best in cases where not much information about the room is available or where a blind estimation is required. RFP, however, requires information about the modes in the room such as the total number of modes to beforehand produce a coherent result. On the other hand, due to the compact processing steps, RFP takes much less processing time than SOMP and hence can potentially be beneficial in certain application such as when online estimation is required or real-time sound field control is involved. In terms of precision, it is important to notice that under sufficient conditions, both methods are capable of producing good estimations for the modal parameters of the room.

5.1.2 Local interpolation

The previous section has analyzed the process of modal identification as well as essential aspects regarding accuracy and robustness. From the outcomes of modal identification, as mentioned in Section 4.2.2, a universal spherical sampling of wave vectors is performed followed by a least square estimation to find the expansion coefficient $C_{n,r}$ which are the last required parameters for the reconstruction. From the outcomes of the entire algorithm, it is now possible to process and interpolate the responses in the room at any point within the geometry. For this numerical studies, the metric of the responses in the room correspond to the transmissions between the source volume flow rate (in m^3/s) and the sound pressure (in Pa) acquired at the measurement points. An example of the Room Frequency Response (also called Room Transfer Function) can be seen in Figure 5.3 for an arbitrary point away from the walls but also not too close to the geometrical center of the room. The interpolation was processed using both the SOMP and RFP method with the same set of 25 microphone positions in the room. It can be seen that both methods can produce an accurate interpolation of the response at this particular point, which strengthen the claim in the previous section.

To further demonstrate the difference between RFP and SOMP in section 5.1.1, one example of the interpolation is plotted in Figure 5.4 where the total number of modes is underestimated.

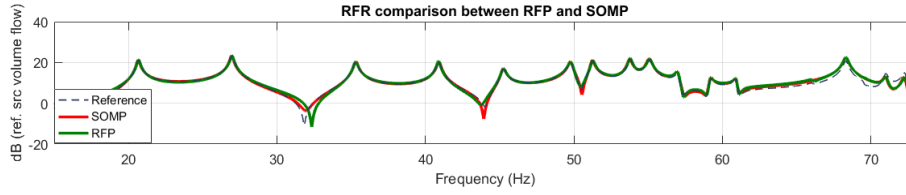


Figure 5.3 – Reconstruction of the RFR at a point inside the room using SOMP and RFP in comparison with the FEM ground truth reference.

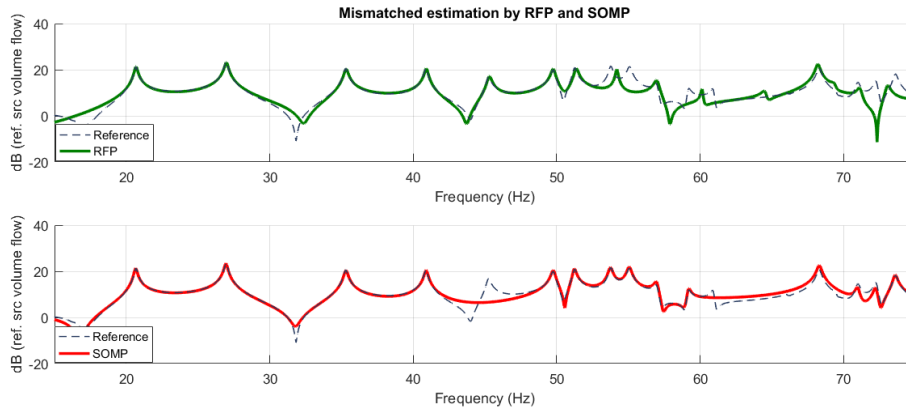


Figure 5.4 – Comparison between the RFRs interpolated with RFP (top) and with SOMP (bottom), when underestimating the number of modes, at a given virtual microphone position. The black curve represents the reference given by FEM simulation

Being an iterative process, SOMP would naturally give a good estimation of all the modes except the ones it does not find. On the other hand, the global curve fitting method using RFP would instead add inaccurate modes that are not from the real system and hence may lead to higher errors. Nevertheless, as explained in Section 5.1.1, testing a few different trials for RFP can certainly solve this problem and hence this method should not be overlooked as its computation time is respectively short and can therefore be advantageous in numerous scenarios.

Coming back to the individual interpolation itself, it must be noted, however, that the high level of precision seen in Figure 5.3 from both method is not guaranteed for every interpolated point in the room and cannot be representative of the reconstruction result as a whole. The recovery error can possibly be higher depending on the location of the evaluated point with respect to the room as well as on the precision of the modal identification results. This, once again, highlights the necessity for a more representative visualization of the sound field to confirm the global validity of the algorithm.

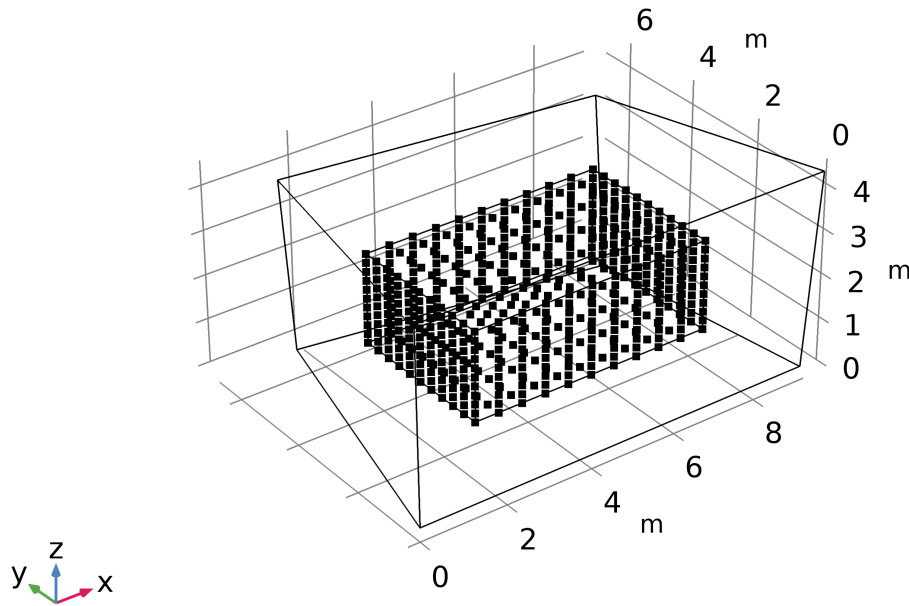


Figure 5.5 – A series of 11x11 microphone grids that is used to sample each of the 6 faces of the inner rectangular evaluation area.

5.1.3 Sound field reconstruction

As discussed in the previous section, an individual interpolation of a single response in the room does not tell much about the overall performance of the reconstruction framework. In this section, the reconstruction process is extended to a large number of locations inside the room to acquire a series of processed time responses of the room. The RFRs of the room can then be produced through the Fourier transform of these time responses. These resulting RFRs open doors to the global reconstruction of the spatial responses of the room at any given frequency of interest.

It is known that, for a room with non-ideally rigid walls ($\alpha > 0$), the mutual orthogonality of mode shapes functions are less reliable near the walls [65, 61]. Furthermore, the recovery of the responses for positions close to the walls approach that of an extrapolation problem rather than our targeted interpolation process. Consequently, the following sound field reconstruction is performed for a predefined shoe-box volume inside the room with each face being at least 1m away from the walls of the room. It is then possible to compare these results with frequency domain simulations, obtained using COMSOL which is considered as the ground truth.

The first sound field reconstruction is performed on the faces of the shoe-box volume. In

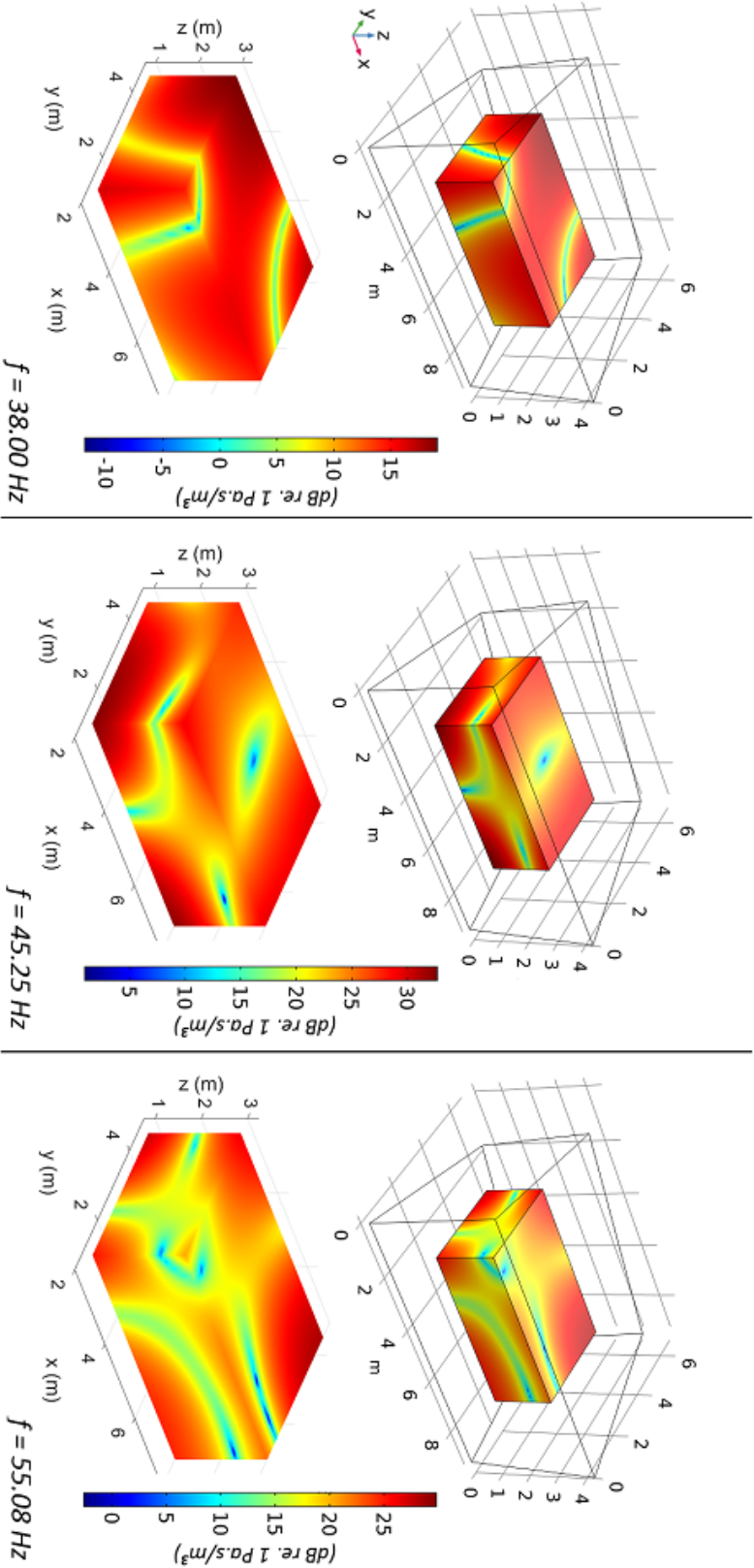


Figure 5.6 – Sound field reconstruction (bottom) at different frequencies for a rectangular area inside the room in comparison with the referencing sound fields from numerical simulation (top).

order to achieve a fine resolution for the reconstruction without putting too much burden on the reconstruction algorithm, each of the faces are made up of a 11×11 equidistant grid of points for reconstruction. This grid can be examined in Figure 5.5. With the reconstructed responses at these 121 discrete points, a simple built-in graphic interpolation algorithm from Matlab can produce a smooth heat map for the spatial responses. With this dense grid of reconstructed responses, it is guaranteed to be able to correctly express the spatial acoustic responses correctly at the very least up to 100 Hz. Figure 5.6 presents three examples of the sound field reconstruction using 25 input microphones placed randomly in the room at three different frequencies at very high spatial resolution. These responses can be directly compared with the reference sound field simulated by the very same numerical analysis included in the same figure.

It can be seen from the comparison that the first reconstruction of the sound field yields highly accurate results for all the three examined frequencies. The existence of the mode shapes is clearly visible in all of the three examples. It is worth noticing that these spatial responses once again proves that compared to a regular shoe-box room, a non-rectangular room possesses numerous room modes that are very difficult to predict. This further shows that the spherical sampling technique for wave vectors is a powerful tool for rooms with complex geometries. Furthermore, the obtained high level of accuracy is maintained in every direction of the 3D depiction. The mode shape behavior are kept consistent throughout all the faces of the rectangular domain. This is thanks to the random placement of the microphones grid in the room as this placement method does not favor any particular shape or directions. A few initial trials using a regularly spaced grid of microphones have been done but none has been able to achieve such global precision in the results. The same can be said about strategies illustrated in 2.6, where the results are not satisfactory with low accuracy and low resolution of the reconstruction. This, once again, emphasizes the advantage of the much-recommended randomness that has been used in other common sparse and low-rank approximation frameworks [97, 70, 29, 57].

It is also worth noticing that even though there might exist some differences when comparing the local sound pressure point by point, the overall shapes and especially the separation between domains of high and low sound pressure are nevertheless precisely depicted. The reconstruction in Figure 5.6 are performed at two distinct exact eigenfrequencies of the room (45.25 Hz and 55.08 Hz) to emphasize the accuracy of the mode shape approximation techniques. On top of that, we also select one random frequency at 38 Hz which lies between two consecutive modes of the room (at around 35 Hz and 40 Hz, respectively) to show that even for a transition frequency in between eigenmodes, the depiction of the sound field still can maintain its precision.

Correlation analysis

Even though it is possible to visually analyze the recovered responses of the room, it remains that we still need to have certain metrics to quantify this accuracy. The normalized Pearson

correlation coefficient for the amplitude of the frequency responses, calculated as below:

$$COR_{\%} = 100 \frac{|\langle |S_f|, |\tilde{S}_f| \rangle|}{\|S_f\| \|\tilde{S}_f\|} \quad (5.2)$$

can be used to evaluate the overall accuracy of the reconstructed frequency response \tilde{S}_f with respect to the reference response S_f . In order to assess the performance of the reconstruction on a larger scale, instead of visualizing just the 2D faces, we proceed to reconstruct the entire rectangular volume of interest. This can be done by performing the framework on a regular 3d grid of 11x11x11 points that samples this rectangular volume. The grid can be seen in Figure 5.7.

A universal evaluation of the $COR_{\%}$ values for these 1331 points yields an average correlation of 99.3% with a standard deviation of 0.8%. Notice that for this evaluation, both the average and the standard deviation value of $COR_{\%}$ is important. The average value explained how accurate the overall estimation of the sound pressure is. On the other hand, the standard deviation has an interesting implication on how much the accuracy differs across the 1331 points. Both value are hence, equally important to assess the precision of the reconstruction framework.

One aspect that should attract our attention is that although $COR_{\%}$ provides a good indication of the overall fitting of the reconstructed signals for the entire concerned bandwidth, it does not provide much information to analyze the precision of the reconstruction frequency-wise. From a single general value of $COR_{\%}$, it is difficult to tell how the accuracy was spread across the bandwidth. A global error evaluation will be introduced later to address this subject.

Reconstruction for rooms with higher damping

Up to this point, the analysis has shown good results for the sound field reconstruction of a lightly damped room with $\alpha = 0.01$. In order to further assess its robustness in more conventional situations with acoustic treatments, the algorithm can be challenged with various room absorption condition. To verify this, the overall absorption coefficient of the walls in the room is first increased to 0.1 and then further increased to a value of 0.3 which can be considered as quite a high absorption properties for the low frequencies of a room. Using the same number and locations of microphones, the reconstruction of the sound field for these two cases is performed using the same procedure as the preceding case. The only difference is due to the expected increase in the damping of the modes, the range of damping value that is used in the modal identification step should be changed accordingly.

Figure 5.8 presents the comparison of the reconstructed sound field for the same mode of the room. In this case, due to an increase in room damping, a same-frequency comparison does

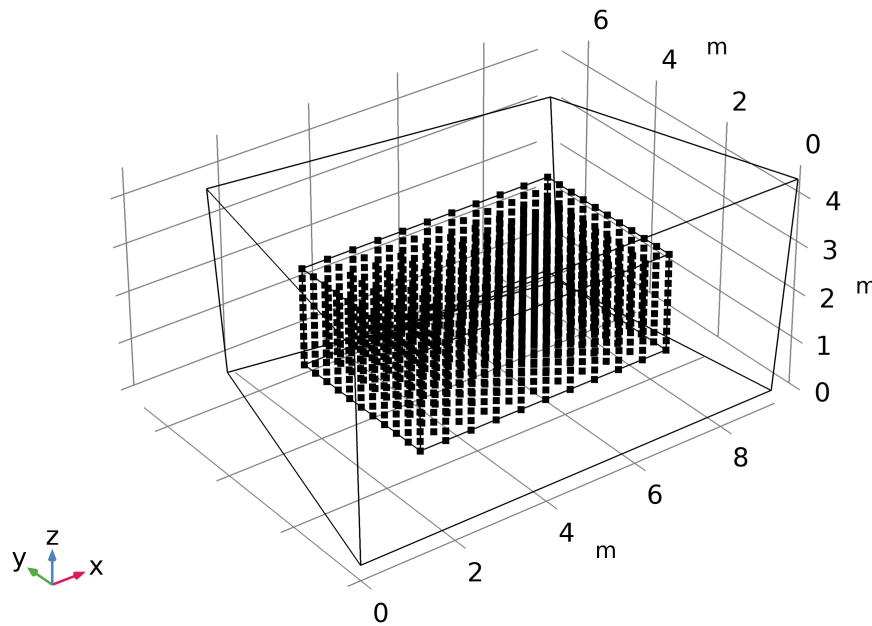


Figure 5.7 – An evaluation 3D grid for the evaluation of the overall accuracy of the reconstructed responses

not make sense anymore, as the physics of the room has been changed aggressively. This, in turn, slightly changes the frequency of each and every modes of the room. This explains why a comparison of the same mode in the room is more logical in this case. From Fig 5.8, it can be seen that the reconstruction results for all these cases still maintain a good agreement with the reference sound field. In terms of the acoustic energy in the room, the framework captures correctly a reduction in the overall sound pressure level in the room as the damping increases. It is interesting to compare the sound field at the same mode in this case as it is possible to observe how the modal responses change for this mode when the damping is increased. The framework once again captures correctly the behavior of the room mode in each cases. Furthermore, the framework also succeeds in rendering the smoothing effect between the node and anti-node of the mode as the room becomes more damped. Finally, it can also be seen that the framework correctly depicts a significant decrease in the dynamic of the sound pressure in the room as the damping increase.

Frequency-wise evaluations

We have previously mentioned some aspects in the reconstruction accuracy that the correlation $COR_{\%}$ cannot express. In order to have a different view point regarding the frequency-wise

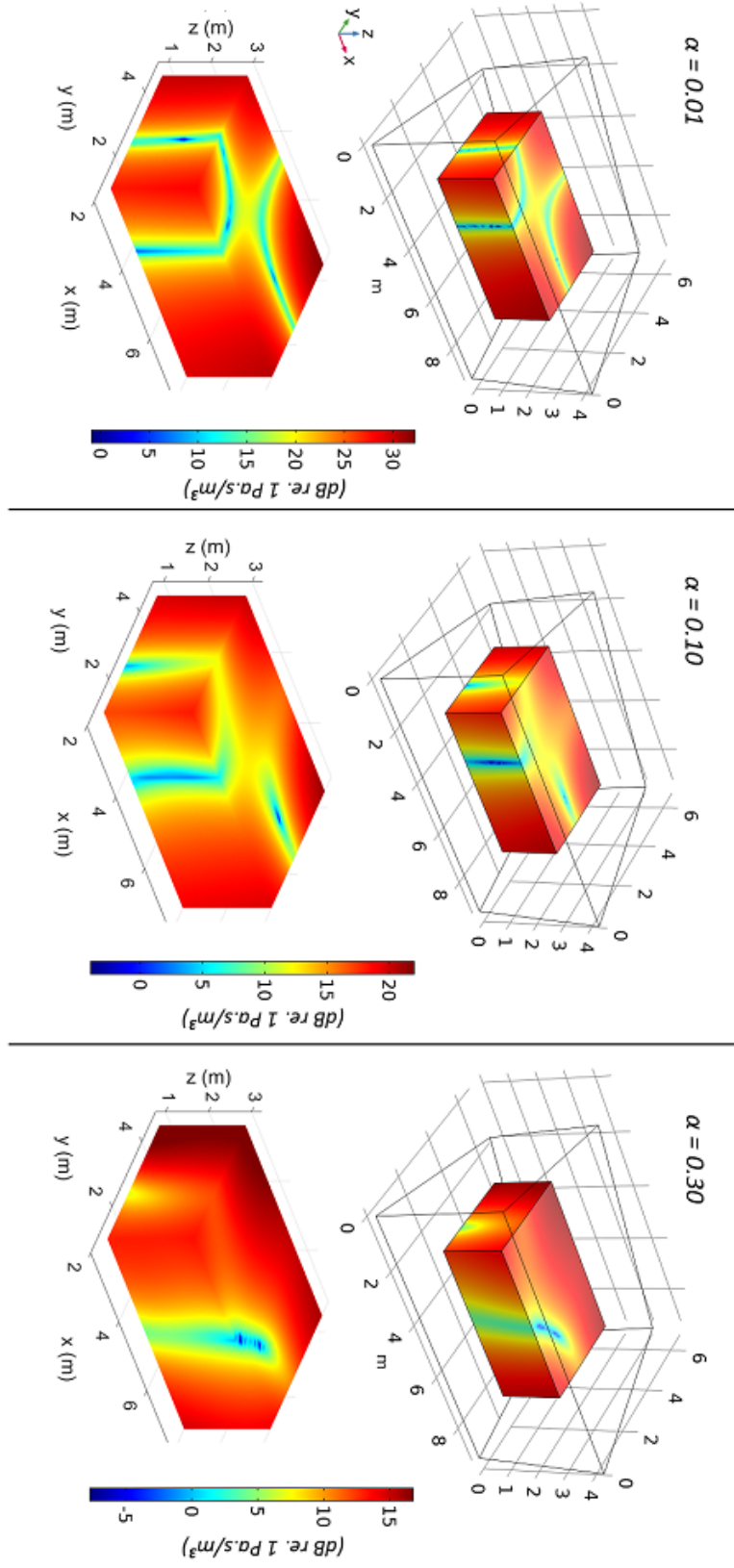


Figure 5.8 – Sound field reconstruction (bottom) compared to the reference (top) for different cases of wall damping ($\alpha = 0.01$, 0.10 and 0.30 from left to right) at the eigenfrequency around 35.3 Hz

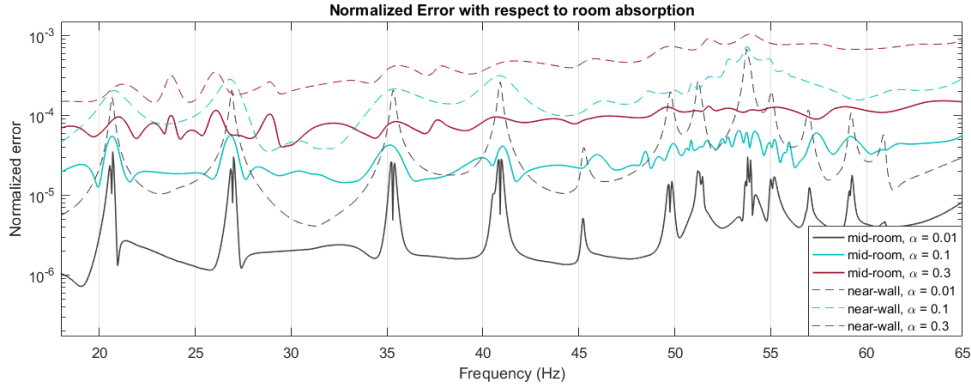


Figure 5.9 – Comparison of the normalized error between different cases of room absorption for the reconstruction of the rectangular volume and a plane near a wall of the room

accuracy of the reconstruction, a dimensionless normalized error is defined as:

$$\overline{e(f)} = \frac{\overline{|S(f) - \tilde{S}(f)|}}{\frac{\int S(f) df}{\Delta f}} \quad (5.3)$$

which specifies the relative error of the reconstruction result at any frequency, normalized by $\int S(f) df / \Delta f$ to discard the dependence on the acoustic energy in the room between different room absorption conditions. This quantity is suitable for comparing the performance of the reconstruction as a function of the room damping because it already accounts for the acoustic energy not absorbed by the room, which, for the same reference source, decreases as the room gets more damped. The non-normalized error calculated as $|S(f) - \tilde{S}(f)|$ can be computed for every point in the 11x11x11 points grid that spatially samples the aforementioned shoe-box test volume. The mean error $\overline{|S(f) - \tilde{S}(f)|}$ for each frequency can then be calculated by averaging the error within this spatial grid. Hence, $\overline{e(f)}$ can be considered as an average spatial reconstruction error metric for each frequency of the room. Figure 5.9 plots this value along the frequency axis for each targeted damping conditions of the room.

It can be seen that there is a decrease in terms of accuracy as the absorption of the walls increases. This is explainable as the orthogonality assumption of the mode shape functions in Eq. (4.3) becomes weaker with higher damping in the room. Furthermore, in overall, the modal identification of the input RFRs is typically more challenging in a room with high damping than in a lightly damped room. Examining across the frequency range, it can be seen that the error generally stays within the same bound and does not fluctuate too much across the frequency axis, although there is tendency of a slight increase of error for higher frequency which will be discussed later on.

One interesting aspect of the overall error curve along the frequency axis is that the normalized error tends to be higher near a room mode. In order to understand this, we should take into account the fact that the acquired RFRs have a certain frequency resolution, this means that there are usually situations where the absolute correct frequency of the room mode lies in between two sample points and hence this not only affects the accuracy of the identification process but also might bring some mismatches when comparing at a single frequency. At low damping conditions where the room modes have high Q-factor, this mismatch is more obvious than in case of higher damping. As can be seen in Fig. 5.9, even at the eigenfrequencies these errors are still less than 10^{-4} for the lightly damped case and hence should not raise much concerns. Nevertheless, in case where it is of great importance to locate to a higher extend the exact frequency of a mode, sampling strategies suggested in [33, 104, 14] can be recommended.

Regarding the overall correlation of the reconstruction, the average $COR_{\%}$ are still favorably high at 98.3 % with 1.8% of standard deviation ($\alpha = 0.1$) and 98.1% in average with 2.1% of standard deviation ($\alpha = 0.3$).

In addition, Figure 5.9 also shows the error of the reconstruction sound field for a plane very close to one of the walls (maximum distance from the wall is 0.1m). It can be seen that the sound field reconstruction close to the boundary of the room induces higher errors, which is anticipated. This is because of the aforementioned non-orthogonality of the mode shape functions along with errors induced by extrapolation instead of interpolation as the concerned reconstructed points are mostly outside of the input microphones domain. Further investigation in the future could shed more light into the contributions of these factors to the overall errors.

The results from this evaluation are especially meaningful in the field of modal equalization. It proves that this particular sound field reconstruction framework can be effectively used to assess the sound field within a room both before and after a given equalization method has been applied, which paves the way for a new available tool in assessing the in situ performance of low-frequency room modes treatments both in frequency and space.

The error curves in Fig.5.9 also shows that the reconstruction error is slightly increased for higher frequencies. From the framework design point of view, there does not seem to be any particularly favor given to the lower-order modes over the higher ones. This indicates that there exists a hidden factor over the parameters estimation procedure that affects the accuracy level. The first possible answer concerns the complexity of the mode shape function. For a room with complex geometries, the complexity of the mode shape functions will undoubtedly increase for higher-order room modes which will, in turn, require a higher number of plane waves to achieve convergence. Even when using the least norm method in Section 4.2.2 to increase the number of plane waves, the compromise between regularization and instability of the mode shape approximation ([24],[4]) often means that the precision still depends heavily on the number of available measurement points. Another possible explanation is that the modal density increases as the frequency gets higher. This means that the average distance (in

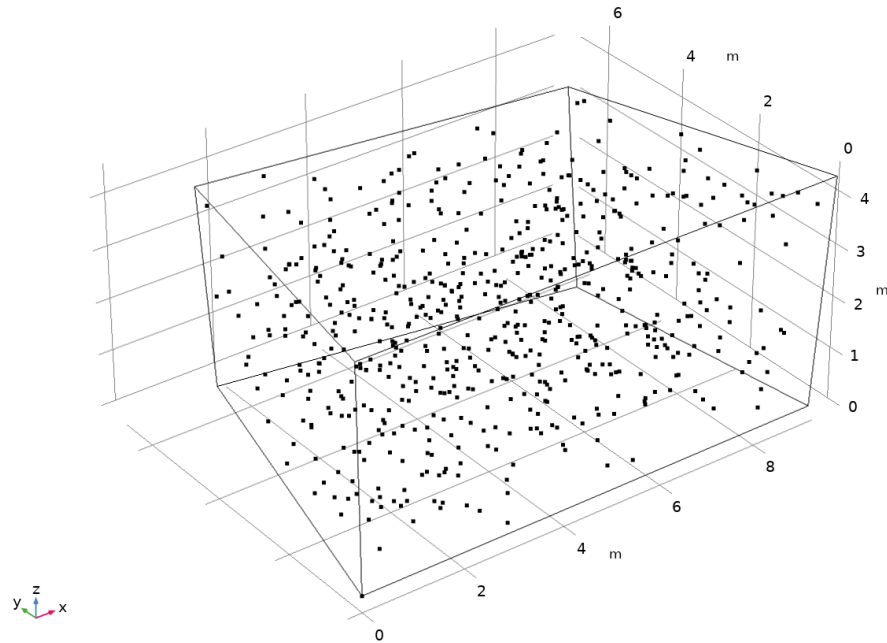


Figure 5.10 – A randomly redundant set of 600 measurement points in the room that can be used to analyze the dependent of the results on the number of input measurements.

Hz) between any two consecutive modes will become smaller and induce more difficulties for the modal estimation process.

Accuracy with respect to number of measurements

Up to now, the number of measurement points (microphones) has not been discussed. As can be seen in Chapter 4, the number of measurements can directly affect the accuracy of both the modal identification and the mode shape approximation process. First of all, a high number of measurement points can ensure that the measurement matrix can capture fully the information on the modal properties of the room, especially in this case, where the number of microphones can be spread randomly in the room. With a smaller number of microphones, it is more likely that the final placements might miss some domains of the room and hence the accuracy of the reconstruction results is more prone to fluctuation space-wise. This is why when evaluating the performance with regards to the number of input measurements, it is important to note that not only the overall (average) accuracy matters, but also the space-wise fluctuation of the results.

In this Section, we propose to analyze the accuracy of the reconstruction results based on the aforementioned Pearson Correlation criteria. The reconstruction can be analyzed based on both the space-wise average as well as the space-wise standard deviation of $COR\%$. Figure 5.12

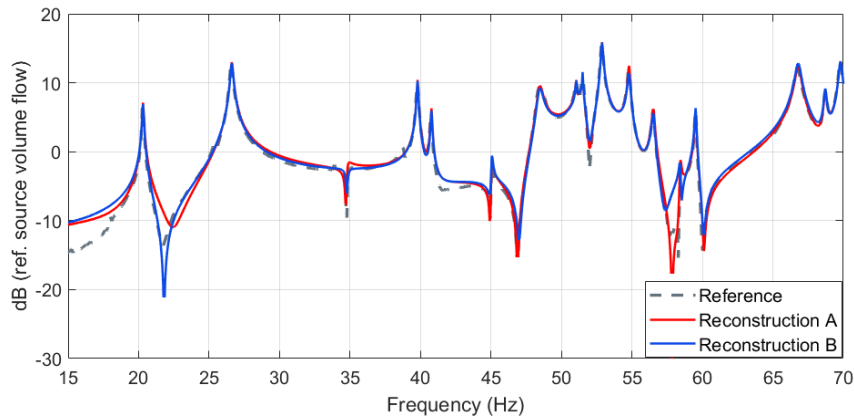


Figure 5.11 – An example of the reconstruction of the RFR at the same evaluation point using two different placements of 15 microphones

shows the comparison of the Pearson Correlation criteria (space-wise average and standard deviation) processed for different numbers of input microphones, and for different absorption coefficients of the walls. For each of the cases, the reconstruction algorithm is repeated multiple times using the same number of measurement points but each time the locations of the input measurements are chosen randomly from a set of 600 randomly spread points in the FEM room model. Figure 5.10 shows the random placement of these 600 measurement points in the room. This procedure is chosen so as to eliminate the bias that could emanate from the placement of the microphones, especially in the cases where the number of microphones is considerably low. To explain this possible bias, Figure 5.11 shows one example where the response of the same point in the room is recovered using two different sets of 15 input microphones.

For each case, the Pearson Correlation is calculated for the same 11x11x11 grid that has been used in the previous sections to sample the shoe-box reconstruction region. It can be observed from Fig. 5.12 that the correlation value gets higher as the number of microphones increases. The standard deviation value mentioned in this figure specifies the standard deviation of the correlation value between different interpolating points in the rectangular reconstruction region. A high standard deviation value will then indicate a highly uneven reconstruction accuracy in which the correlation of the reconstruction varies significantly depending on the location of the interpolation. This could, in turn, severely affect the accuracy of the reconstructed mode shapes behavior in the room. Conversely, a low standard deviation indicates that the spatial reconstruction result is stable and can be trusted.

Figure 5.12 shows that the average correlation values improve as the number of measurement points increases. Moreover, the standard deviation value also decreases significantly when more measurement points are used for the framework. This indicates that while the overall reconstruction gets more accurate, the estimation accuracy also becomes uniformly more

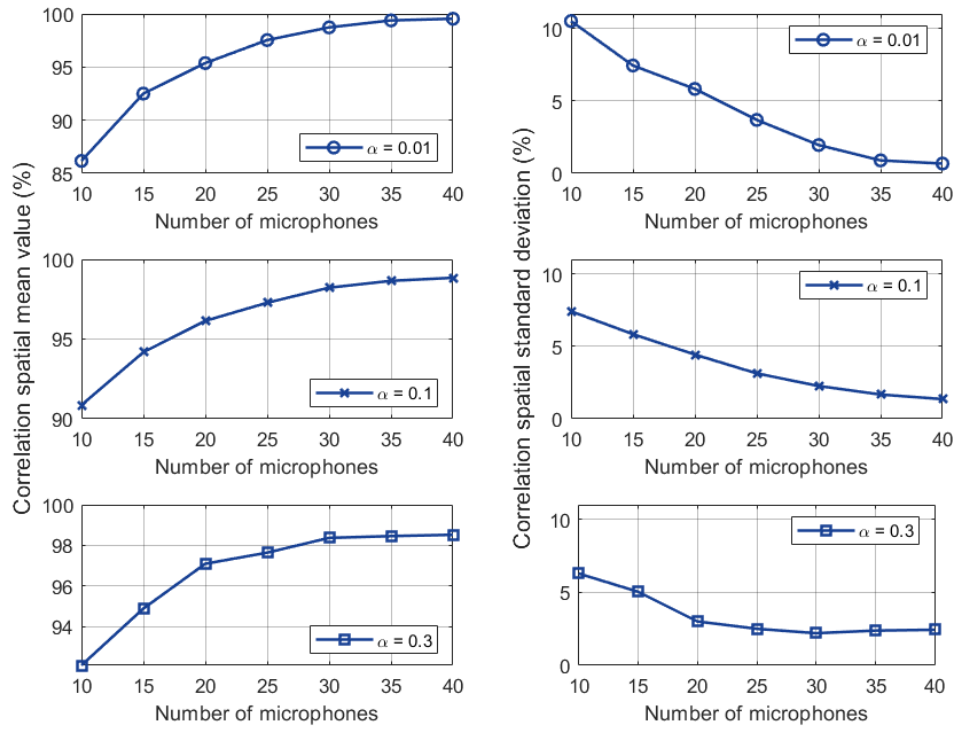


Figure 5.12 – Analysis of the Pearson Correlation of the reconstructed sound field with respect to the number of measurement points.

stable across all interpolation points in the evaluation grid. One of the reason, as mentioned before, is that more measurement points allows a better the chance at estimating correctly the room modes information. A higher number of microphones also allows a more reliable mode shape approximation using a higher number of sampled wave vector. It can also be observed that for the analysis within a fixed bandwidth, the performance typically becomes stable and reliable when a certain number of measurement points is reached. In the case of a lightly damped room, for instance, a grid of size 1331 within a volume of 40 m^3 can be reconstructed with a high accuracy of 98.5% using just 30 input measurement points which is an effective result for a practical number of microphones. Furthermore, even with only 20 microphones, the result is still considered stable with a trusted average correlation around 95%.

5.2 Experimental results

Thanks to the analysis in Section 5.1, numerous aspects of the reconstruction framework have been tackled. These analysis provide the foundation for the experimental validation of the framework as a whole. In this section, the same reconstruction framework can now be applied

to actual measurements inside the reverberation chamber at EPFL (Figure 5.13), which has the same geometry as the FEM model in Section 5.1. The reflective panels in the actual room are expected to be quite acoustically transparent to the measurements of interest as the main focus is on the very low frequency range (<100 Hz where $\lambda > 3\text{m}$).

The source is a custom-made subwoofer in a closed wooden cabinet. The subwoofer is placed near a corner of the room in order to be able to excite all room modes at low frequencies. The microphones (PCB 378B02 1/2" microphones) are spread randomly in the room to replicate the previous numerical analysis. The placement strategies can be quickly pre-examined using the numerical simulation first to guarantee a stable reconstruction. The location of microphones are measured with an uncertainty of less than 1 cm which, when compared to the room sizes is negligible. The reference for the acoustic response in the room is chosen to be the velocity of the source. This velocity can be measured with a laser velocimeter (Polytec OFV 500) placed in front of the loudspeaker diaphragm.

In order to evaluate the reconstruction results, two main methods are proposed. One method is to directly compare the reconstructed sound field to the simulated one in FEM. This method can be very useful in verifying the faithfulness of the spatial rendering aspect of the reconstruction results. However, this method is not recommended for a point-by-point comparison as it is difficult and cumbersome to meticulously match the FEM model with the real one, since it relies heavily on the absorbing properties of the room which are not accurately known and conventional methods are prone to high error at low frequencies [38, 103, 73, 26]. Moreover, the reference used for processing the RFRs can be different between the simulation (volume flow) and the actual case (velocity) and it is difficult to accurately match the physics of the source as well as its position. This is why, besides this method, a small part of the available points can be reserved to serve as an evaluation set. Combining these two evaluation methods provides a more concrete analysis of the reconstruction framework using the experimental data.

In this experiment, signals from 25 different microphone locations are used as the inputs of the algorithm to render the sound field up to 75 Hz (within which about 20 modes can be observed, with the first mode of the room located at around 20 Hz). The microphones are spread randomly in the room but are chosen so that they are distributed evenly space-wise to practically cover the area of the reconstructed rectangular volume. Figure 5.14 visualizes the spatial comparison between the reconstructed sound field and the reference one obtained from numerical simulation at the same eigenmodes. As previously mentioned, due to the fact that the numerical model cannot be perfectly matched with the real room, there will certainly be some differences in terms of the exact frequency of the eigenmodes.

Comparing the results in Fig. 5.14, it can be observed that similar to the numerical results in the previous section, the reconstructed sound fields from real measurements yield highly accurate spatial recovery. The mode shapes are visible and the locations of nodal lines are correctly depicted with high spatial resolution. The first of the two presented mode, at 35.20



Figure 5.13 – Measurement set up in a real reverberation chamber in the laboratory

Hz, can be refer to as being very close to a traditional (1,1,0) tangential mode (2D mode) whereas the second one at 51.25 Hz presents a much more complex modal behavior due to the complex geometries composed of angled walls of the reverberation chamber. Nevertheless, the reconstructed framework can be seen to correctly render both of the modes with equally high accuracy. Once again, small mismatches at a few discrete points are even more likely to be expected, knowing that there are inevitable differences between the real and simulated room. However, the most important thing is that the overall spatial representation remains to be faithful.

Using an evaluation set of 30 other microphone locations within the domain of interest, the results also agree with the previous simulation validation. For the reconstruction that uses SOMP as a modal estimation method, the average correlation stays at 97.8 % with a value of 1.89 % for standard deviation. As expected, this evaluation result is slightly less accurate than the average correlation obtained with simulations but can still be classified as highly reliable. As mentioned earlier in Section 5.1.1, SOMP is particularly robust and tends to perform well even when a priori information is not readily available. On the other hand, under the same circumstance, running the reconstruction framework with a semi-supervised RFP global curve fitting produces a slightly lower average correlation of 96.9 % with a higher standard deviation of 2.3%. This result also agrees with the analysis in Section 5.1.1 regarding the different in the nature of the two modal identification methods. Three discrete examples of the reconstructed RFRs by both RFP and SOMP are plotted in Figure 5.15 and compared to the actual measurements from the evaluation set. Generally, without a detailed supervision and calibration, the RFP method will return a slightly less accurate result than SOMP as can be observed from the figure. However, its processing speed is much faster and hence could allow

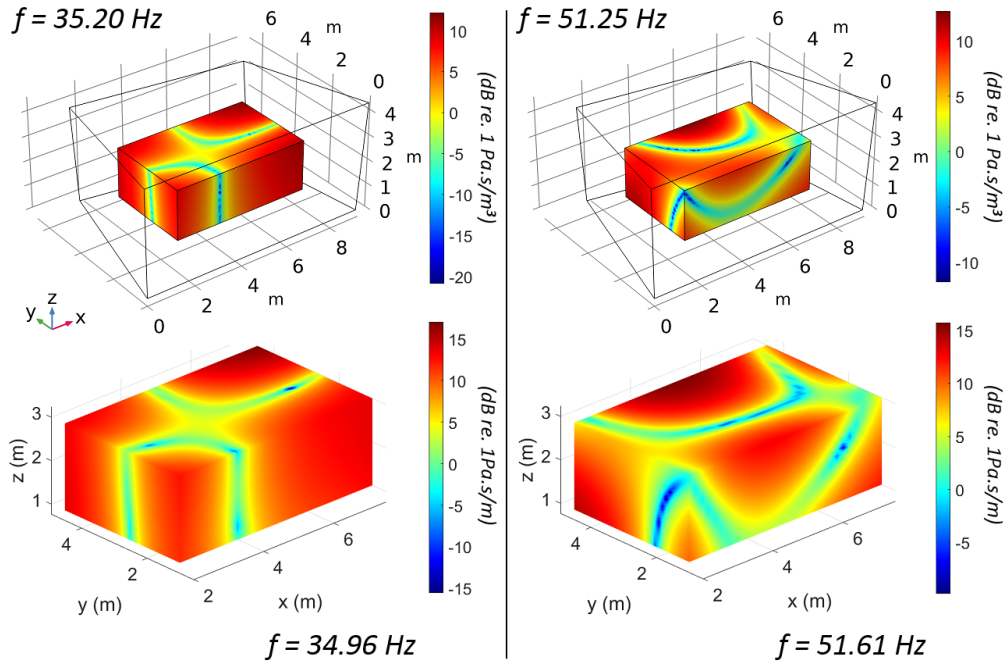


Figure 5.14 – Sound field reconstruction from real measurements (bottom) at two distinct eigenmodes (left: around 35 Hz, right: around 51 Hz) as compared to the same ones from simulations (top)

for quick re-calibration depending on the situation.

It should be noted that in general practice, when the room geometries and wall absorption coefficient are not specifically known, evaluation set like this along with the comparison parameters such as the correlation values are among the few available indications to know whether the reconstruction results are reliable. Therefore, practically, it is advised to always have a reserved evaluation set inside the domain of interest to navigate the adequate number of microphones required for any certain objective. A promising sign for the practical use of the framework is that the standard deviation for the correlation value remains to be relatively small for both analyzed cases. This indicates that the size of the evaluation set of microphones does not have to be large due to the fact that the correlation value does not fluctuate much between the evaluation microphones.

Regarding the experiment set-up, as the number of microphones is practically small, a blind random placement of microphones might leave out crucial areas of the room. Hence out of all the possible randomization, it is advised to choose an appropriate placement that does not leave out crucial areas of the region of interest. Moreover, placement technique like the one suggested in [24] might also be used to improve the recovery results. Lastly, it should be noticed that the measurement was conducted in a reverberation chamber without removing the reflective diffusing panels (Figure 5.13). This shows that the framework is robust enough to perform well even in a practical non-empty room.

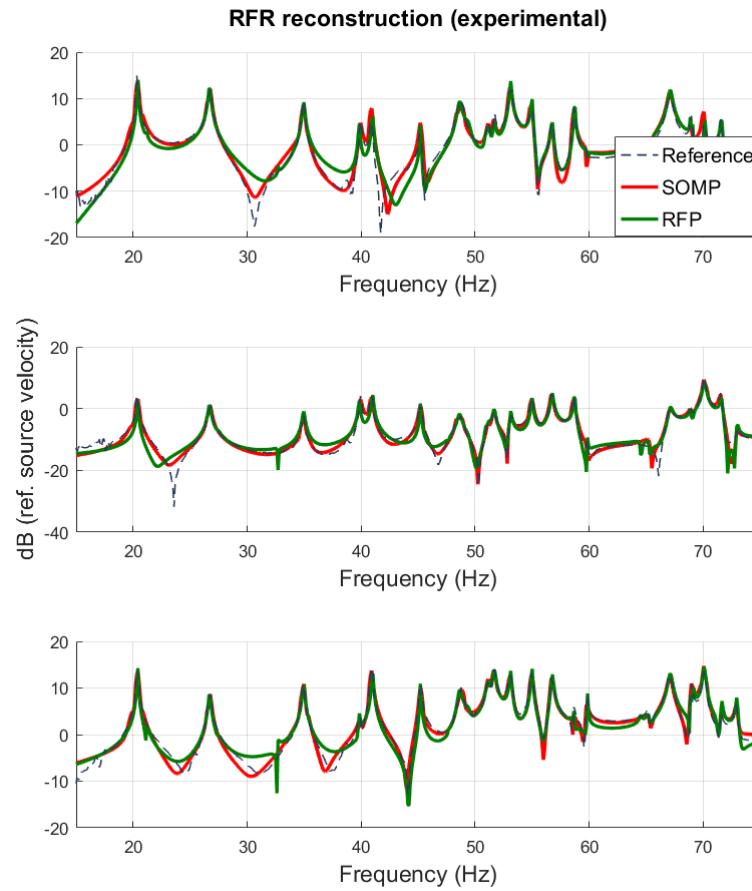


Figure 5.15 – RFR reconstruction from 25 measurements for three different evaluation points in the room using RFP and SOMP (correlation ranging between 97% and 99%).

5.3 Conclusion

In this chapter, the sound field reconstruction framework in Chapter 4 has been tested both with simulation data and real experiment in a reverberation chamber. The reconstruction was first analyzed in terms of its modal estimation performance. The two proposed methods of Simultaneous Orthogonal Matching Pursuit (SOMP) and global curve fitting using Rational Fraction Polynomials (RFP) are both capable of obtaining the modal parameters of the room. Out of the two, SOMP has been shown to require little to no supervision while RFP requires more of a semi-supervised trial and error methods. After the modal parameters are acquired, the framework can reconstruct the sound field in a chosen area in the room. The reconstructed sub-space was chosen to be a large rectangular volume within the reverberation chamber. Using the FEM simulation inside a reverberation chamber model, it has been observed that the reconstruction is able to achieve highly accurate results with the correlation percentage higher than 95 %, even closer to 99 % in many cases. The sound field and its accompanying

mode shapes behavior are also precisely recovered through the framework. The reconstruction framework have also then been assessed in terms of multiple revolving aspects such as room damping, number of microphones, spatial errors and have shown a highly reliable and robust performances. After being validated by the simulation, the same procedure has been applied to practical experiments in a real reverberation chamber at EPFL. The reconstruction framework once again achieved highly accurate results for not just individual response recovery but also for the entire sound field of the reconstructed rectangular volume.

From the results of this chapter, it becomes clear that the reconstruction framework can perform with high level of accuracy and resolution for both the un-damped and damped situation of the room. This means that the reconstruction framework could potentially be used as an assessment tool to analyze the spatial performance of not just passive but also active sound absorption method in the room. The next chapter will investigate this application and analyzed its results as well as applicability.

6 Reconstruction of active-controlled sound field

In Chapter 2, we have introduced the room modes phenomena at low frequencies along with its effects on the listening experience. Due to the discrete room modes and their irregularities in both spatial and frequency domain, these effects, which include coloration and masking effects of the perceived sound, are likely to alter and damage the listening experience in the room. In order to reduce the effects of room modes, various techniques have been introduced towards equalizing room modes. Amongst the available methods, passive techniques such as foam based absorbers or bass traps are either ineffective at low frequencies (< 100 Hz) or pose certain inconvenience regarding the geometries and volume occupancy. Because of these reasons, various active equalization methods have been developed to improve the damping of the room at low frequencies. Although these techniques can be vastly diverse, at the end of the day, the assessment methods are usually limited to a few discrete measurements in time or frequency domain to analyze metrics such as individual modal damping or modal decay times.

In Chapter 4, we have developed a novel sound field reconstruction framework which allows the spatial reconstruction of the sound field in a large sub-space of the room. The results of the reconstruction framework have been analyzed with different scenario and proved to be highly accurate and robust even for a non-rectangular room in Chapter 5. This framework further allows for the analysis of the modal behaviors of the room in space.

In this chapter, we aim to bridge the gap between the two aforementioned domains of Room modes equalization and Sound field reconstruction by showing the vast possibilities that Sound field reconstruction can bring to assist towards the equalization of room modes.

6.1 Closed box loudspeaker system and the Electroacoustic absorbers

In this section, we will briefly review the closed box electrodynamic loudspeaker system and extend the analysis towards the introduction of the electroacoustic absorbers and its governing control design. A more detailed analysis on its target impedance will be addressed in Section

7.1.1.

6.1.1 Closed box Electrodynamic loudspeaker system

In order to understand the design theory behind the Electroacoustic absorber, we briefly review the closed-box Electrodynamic loudspeaker system. For a simple closed box electrodynamic loudspeaker system, using the Newton's 2nd law where the sum of forces is equal to the mass multiplied by acceleration, the governing motion equation can be expressed as:

$$S_d P_f(\omega) = \left(Z_{ms}(\omega) + \frac{S_d^2}{j\omega C_{ab}} \right) V(\omega) + BlI(\omega) \quad (6.1)$$

where P_f is the pressure at the front of the diaphragm, V is the velocity of the diaphragm, S_d as the effective piston area of the membrane and Bl as the force factor of the loudspeaker. Z_{ms} is the mechanical impedance of the loudspeaker which can be expressed as:

$$Z_{ms}(\omega) = j\omega M_{ms} + R_{ms} + \frac{1}{j\omega C_{ms}} \quad (6.2)$$

with M_{ms} , C_{ms} and R_{ms} are, respectively, the mass, mechanical compliance and mechanical resistance of the loudspeaker.

The term $\frac{S_d^2}{j\omega C_{ab}} V(\omega)$ demonstrate the force at the back of the membrane due to the loudspeaker enclosure where C_{ab} is the acoustic compliance of the enclosure and can be expressed as:

$$C_{ab} = \frac{V_b}{\rho c^2} \quad (6.3)$$

with V_b as the effective volume of the enclosure, ρ as the density of the medium (air) and c as the sound celerity. With the formulation in Eq. 6.1, we can combine the mechanical impedance and the term responsible for the enclosure volume as:

$$Z_m(\omega) = Z_{ms}(\omega) + \frac{S_d^2}{j\omega C_{ab}} = j\omega M_{ms} + R_{ms} + \frac{1}{j\omega C_{mc}} \quad (6.4)$$

where Z_m is considered to be the total mechanical impedance of the closed box loudspeaker system with C_{mc} as the total mechanical compliance where:

$$C_{mc} = \frac{C_{ms}C_{ab}}{C_{ms}S_d^2 + C_{ab}} \quad (6.5)$$

Finally, the equation in Eq. 6.1 can be shorten into:

$$S_d P_f(\omega) = Z_m V(\omega) + Bl I(\omega) \quad (6.6)$$

6.1.2 Electroacoustic absorber

The control design of the electro-acoustic absorber is based on a feed-forward control using a single sound pressure input from a microphone in front of the membrane of a loudspeaker, which is driven by a controlled current with the objective of achieving the target impedance at the diaphragm of the speaker.

The control diagram of the system can be seen in Fig 6.1 in the s domain of the Laplace transform. The specific impedance in front of the closed-box loudspeaker diaphragm can then be calculated as:

$$Z(s) = \frac{P_f(s)}{V(s)} = \frac{Z_m(s)}{S_d - Bl\Theta(s)} \quad (6.7)$$

Assuming that a target specific acoustic impedance of Z_{st} is imposed in front of the diaphragm, the transfer function of the controller $\Theta(s)$ can be expressed as:

$$\Theta = \frac{I(s)}{P_f(s)} = \frac{S_d Z_{st}(s) - Z_m}{Bl Z_{st}(s)} \quad (6.8)$$

In order to correctly imposed the control transfer function, an accurate evaluation of the closed box loudspeaker mechanical parameters such as S_d , Bl and Z_m and its components is required.

Depending on the application, the design of the controller Θ can be changed produce a target specific acoustic impedance Z_{st} in front of the diaphragm of the electroacoustic absorber. In the most rudimentary case, the target specific acoustic impedance can simply be a real

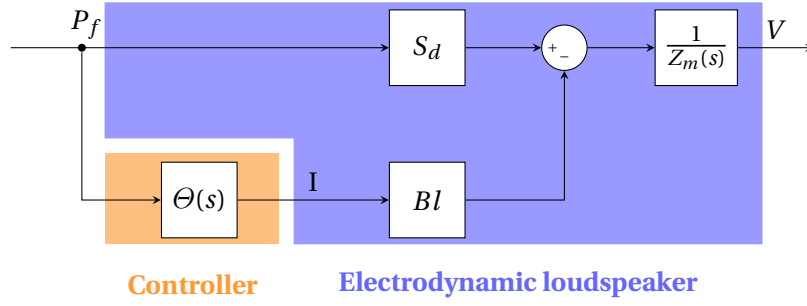


Figure 6.1 – The control design behind the electroacoustic absorber where the input current of the loudspeaker is controlled by the measured total sound pressure at the diaphragm

value, denoted R_{st} which represents a single resistor. However, it has been shown in [16] that to achieve efficient sound absorption and avoid instability, it is more efficient to use a 1-DOF frequency-dependent target impedance with a flexible parametric form of:

$$Z_{st}(s) = s \frac{\mu_M M_{ms}}{S_d} + R_{st} + \frac{\mu_C}{s S_d C_{mc}} \quad (6.9)$$

where $\mu_M \in (0, 1)$ and $\mu_C \in (0, 1)$ are controlled parameters that can be used to extend the bandwidth of the target impedance. The center frequency of the absorption performance can be calculated respectively:

$$f_0 = \frac{1}{2\pi \sqrt{M_{ms} C_{mc}}} \sqrt{\frac{\mu_C}{\mu_M}} \quad (6.10)$$

In this case, using μ_M and μ_C both the center frequency and the bandwidth of the 1-DOF impedance can be altered to match or satisfy certain absorption strategies.

Following the analysis of target impedance for multiple room modes of the room (which will be later discussed in more details in 7.1.1), it is found that for a fixed total effective absorption area of the absorber, each eigenmode of the room provides a different optimal value for the target impedance that can be set in front of the diaphragm of the active absorbers. This makes the 1-DOF absorber less efficient because even though its center frequency and bandwidth could be modified, only one value of resistance R_{st} can be applied in front of the diaphragm. This is why in [100], it is suggested that a multiple degree of freedom resonators model for the target impedance can be applied instead. For instance, for the 3-DOF design that is applied to the electroacoustic absorber available in EPFL, the final impedance is made up of three distinct 1-DOF target impedances in parallel. Each of these impedance is expressed in the

Table 6.1 – Small signal parameters of the closed-box loudspeaker system in the existing Electroacoustic Absorber at EPFL

| Parameter | Notation | Value | Unit |
|-----------------------------|----------|--------|-----------------------------------------------|
| Enclosure volume | V_b | 10 | dm^3 |
| Moving mass | M_{ms} | 14.67 | g |
| Mechanical resistance | R_{ms} | 1.31 | $\text{N} \cdot \text{s} \cdot \text{m}^{-1}$ |
| Total mechanical compliance | C_{mc} | 242.35 | $\mu\text{m} \cdot \text{N}^{-1}$ |
| Effective piston area | S_d | 151 | cm^2 |
| Force factor | Bl | 5.98 | $\text{N} \cdot \text{A}^{-1}$ |
| Air density | ρ | 1.2 | $\text{kg} \cdot \text{m}^{-3}$ |
| Air celerity | c | 343.86 | $\text{m} \cdot \text{s}^{-1}$ |

frequency domain as:

$$Z_{st_k}(\omega, v_{2k-1}, R_{st}, v_{2k}) = j\omega \frac{M_{ms}}{S_d v_{2k-1}} + R_{st} + \frac{1}{j\omega S_d v_{2k} C_{mc}} \quad (6.11)$$

with $k \in [1, 3]$ for the case of a 3-DOF resonators. In this scenario, for each 1-DOF impedance, there are three variable parameters which are v_{2k-1}, R_{st}, v_{2k} to adjust its target resistance, center frequency and bandwidth.

The resulting target 3-DOF impedance which is made up of three parallel 1-DOF impedance can be expressed as:

$$Z_{st_{3-DOF}}(\omega) = \frac{1}{\sum_{k=1}^3 \frac{1}{Z_{st_k}(\omega, v_{2k-1}, R_{st}, v_{2k})}} \quad (6.12)$$

6.1.3 Existing prototype of electroacoustic absorber

In this section, the design and settings of the existing 3-DOF electroacoustic absorber at the facility in EPFL is briefly discussed. The layout of the prototype ([99]) is included in Fig. 6.2

The prototype consists of four Peerless SDS-P830657 loudspeakers in individual enclosure boxes. The overall size of one prototype is (0.3m x 0.3m x 0.62m). Due to technological constraints, the 4 loudspeakers are connected in series which in this case, means that they are driven by the same current. In this prototype, only one microphone is used as the input of the control system and is placed on one vertical side of the enclosure. The parameters of the closed-box loudspeaker are included in Table 6.1.

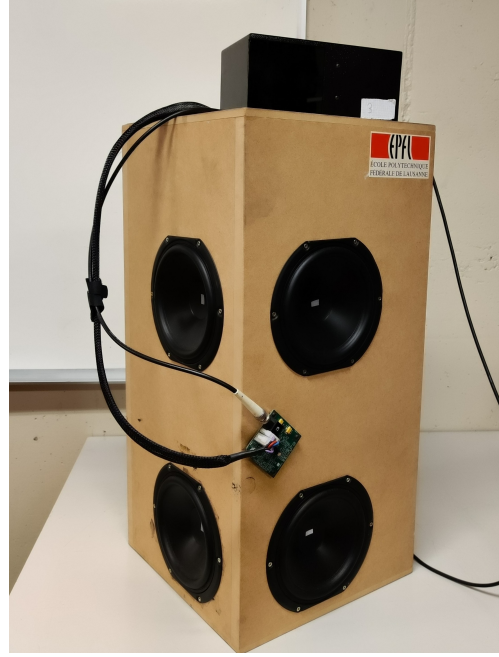


Figure 6.2 – The existing Electroacoustic Absorber at EPFL with 4 current-controlled loud-speaker and one input electret microphone

The settings for this particular 3-DOF absorbers is the same as the one that has been computed in [99]. The target impedance $Z_{st3-DOF}$ is composed of 3 component impedances of Z_{st1} , Z_{st2} , Z_{st3} . The designs of these component 1-DOF resonator are listed in table 6.2. The bode plot of the resulting target specific impedance is plotted in Fig. 6.3. This particular settings for the 3-DOF Electroacoustic Absorber is based on an objective function for optimal resistances along the frequency range of interest. The aim is to minimize the modal decay times of simulated room responses. This will be discussed in more details in Chapter 7 to emphasize on how a sound field reconstruction framework can open door to a different approach to this problem. For now, we will simply use this settings for the numerical studies in the next step.

6.2 Sound field assessment of existing active absorption methods

In this section, the sound field reconstruction framework will be used specifically to assess the resulting sound field of an active room modes damping method, in this case, a multi-degree-of-freedom (n-DOF) electroacoustic absorber developed in EPFL [99] which consists of closed-box controlled loudspeaker systems.

The analysis of the sound field reconstruction framework under different damping conditions in Section 5.1 has laid a concrete foundation for the work in this section. In Section 5.1, it has been shown that the sound field reconstruction framework performs well not just for a lightly damped room but also for room with higher damping. Although the accuracy tends to reduce

Table 6.2 – Current target impedance for the 3-DOF resonator model of the electroacoustic absorbers

| $Z_{st3-DOF}$ | | | | | | | | |
|---------------|------------------------------------------------|-------|------------|------------------------------------------------|-------|------------|------------------------------------------------|-------|
| Z_{st_1} | | | Z_{st_1} | | | Z_{st_1} | | |
| v_1 | R_{st_1} | v_2 | v_3 | R_{st_2} | v_4 | v_5 | R_{st_3} | v_6 |
| (–) | $\text{Pa} \cdot \text{s} \cdot \text{m}^{-1}$ | (–) | (–) | $\text{Pa} \cdot \text{s} \cdot \text{m}^{-1}$ | (–) | (–) | $\text{Pa} \cdot \text{s} \cdot \text{m}^{-1}$ | (–) |
| 1.86 | 85.99 | 2.86 | 1.32 | 54.81 | 1.11 | 3.07 | 57.32 | 21.03 |

as the damping of the room gets higher, the overall reconstruction correlation still remain at a highly reliable value. This suggests that the reconstruction framework can be used to evaluate the performance of active room absorption methods by assessing the rendered sound field of the room with and without the presence of the concerning equalization methods.

6.2.1 Numerical studies

In this section, we analyze the same FEM model of a non-rectangular reverberation chamber used in Section 5.1. Within the numerically generated room, four 3-DOF electroacoustic absorbers are put close the for corners of the room. These 3-DOF absorbers have exactly the same design and also have the same target impedance. Each Electroacoustic Absorber consists of a single microphone and 4 closed box loudspeakers which responses are controlled by the same current. The current sent to the loudspeakers are derived from a feed forward control in Fig.6.1 using the sound pressure measured by a single microphone attached to the wooden enclosure of the absorbers as described in the previous sections. The parameters of the speakers used as absorbers in the model are set based on the real measured parameters in table 6.1. Regarding the reaction with the volume in the boxes, to simplify the design and and processing load on the numerical study, the volume of the boxes were subtracted from the volume of the room. In order to then account for this volume, the value of C_{mc} is included in the mechanical impedance of the speaker instead of C_{ms} (which does not include the compliance produced by the volume $V_b = 10\text{dm}^3$ of the enclosure).

In order to properly design these absorbers in a numerical study, a series of variables are formed within the variables section of COMSOL to simulate the feed forward behavior of the control loop. The input signal of each Electroacoustic Absorber is designed as a single measurement point attached to the absorber's enclosing box. The pressure that is numerically measured from the sensing microphone is used to feed into the control of the four loudspeakers. In order to initiate the loudspeakers based on the control loop, a velocity is applied onto

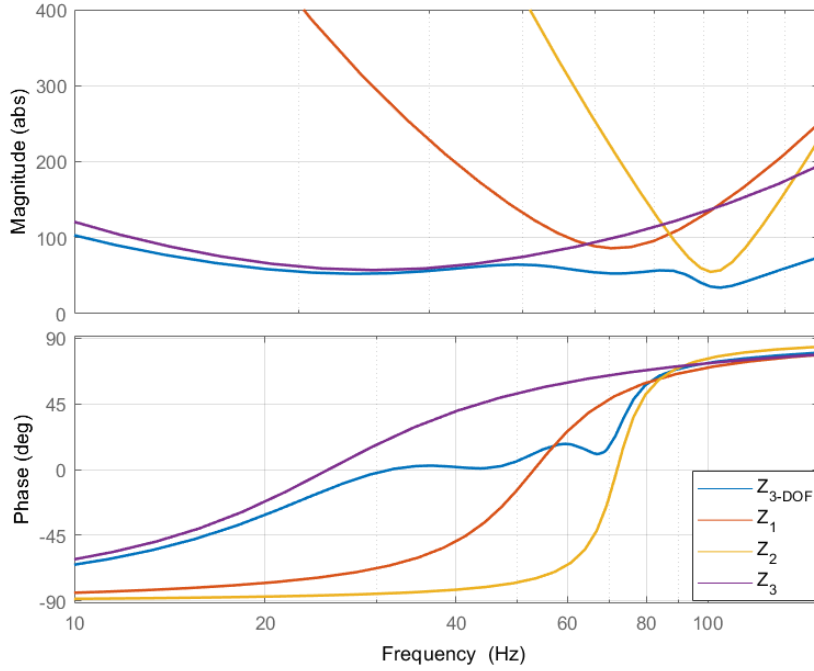


Figure 6.3 – Bode plot of the 3-DOF target impedance in Table 6.2 and its individual 1 dimensional resonators.

the membrane of all four loudspeakers:

$$V(\omega) = \frac{S_d P_f(\omega) - B I I(\omega)}{Z_m(\omega)} \quad (6.13)$$

where the value of $I(\omega)$ can be calculated from the target impedance in Table 6.2 using Eq.6.7. The four absorbers are now ready for the numerical simulation. A brief simulation was then done to quickly check if the simulated absorbers are working properly in the confined room. Figure 6.5 plots the responses at the same point in the room in two scenario, with and without the four absorbers. It can be seen that the absorbers are working properly and exhibit the same behavior as has been described in [99].

After the settings have been applied to the 4 Electroacoustic Absorbers, we repeat the same simulation procedure in Section 5.1, where there are 600 simulated positions of microphones in the room that can be potentially used as the input of the reconstruction framework developed in Chapter 4.

Repeating the same reconstruction procedure detailed in Section 5.1, we achieve the reconstruction parameters of the room, namely the modal parameters ω_n and δ_n as well as a list of

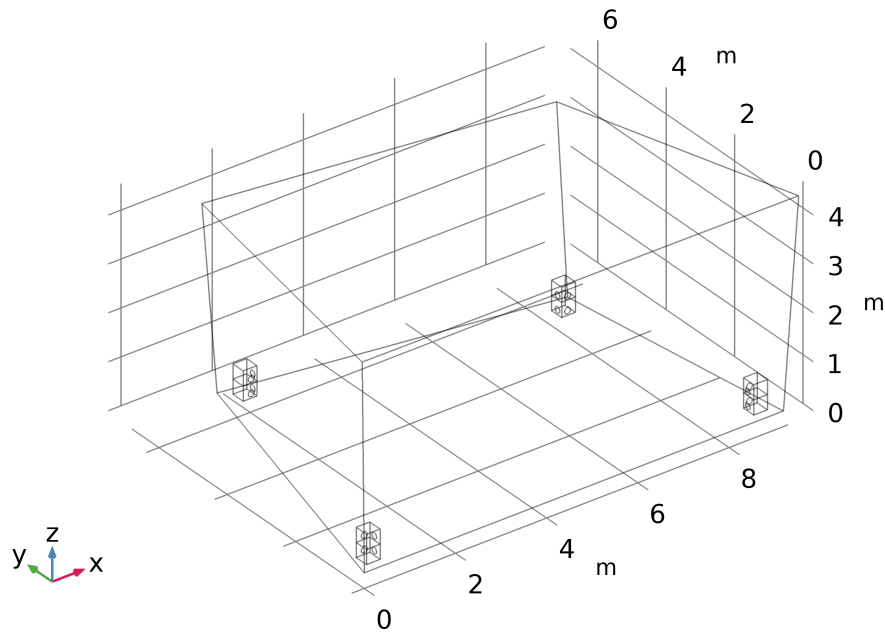


Figure 6.4 – Numerical model of the reverberation chamber with the presence of the Electroacoustic Absorbers at 4 corners of the room

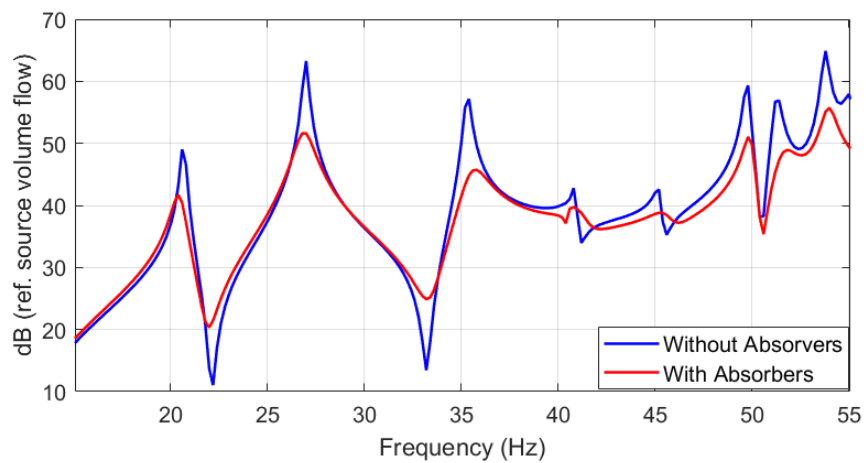


Figure 6.5 – Comparing the response of at a single point in space with and without the presence of the 4 Electroacoustic Absorbers

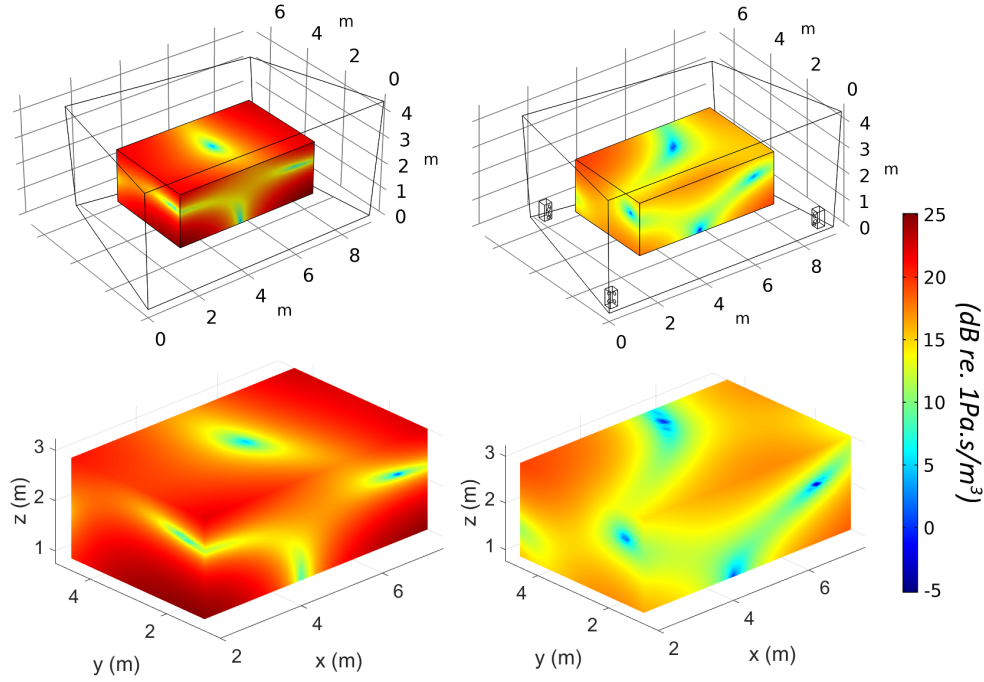


Figure 6.6 – The sound field before (left) and after (right) applying the electroacoustic absorbers are depicted precisely using the reconstruction algorithm (bottom) when compared to the simulation results (top) at $f=45.25$ Hz

$N \times R$ reconstruction plane waves $e^{j\vec{k}_{n,r}}$ and their corresponding expansion coefficients $C_{n,r}$. Using these parameters to reconstruct the sound field over the faces of the same inner rectangular volume described in Fig.5.5 provide us with a visual evaluation of the reconstructed sound field. Figure 6.6 (at $f = 45.25$ Hz) and Figure 6.7 (at $f = 55.07$ Hz) depicts the sound field reconstruction of the room both with the on and off case of the electroacoustic absorbers and comparing both of them to the ground truth sound field results using the simulation from COMSOL.

From the comparison, it is clear that the sound field reconstruction for the room, similar to the results in Chapter 5, achieve highly accurate and reliable results. The most significant aspect of this reconstruction is that it successfully depicts all the changes in the modal behavior of the room while there are complex impedances in front of the 16 membranes. Moreover, the results show that the reconstruction framework performs well even when there is local absorption in the room.

The two example frequencies reflect two different eigenmodes of the room. The eigenmodes are chosen to be compared in this case as the changes in terms of the sound field between the ON and OFF case of the electroacoustic absorbers are more drastic at or near the eigenmodes of the room. As can be seen from the comparison, the reconstructed sound field correctly

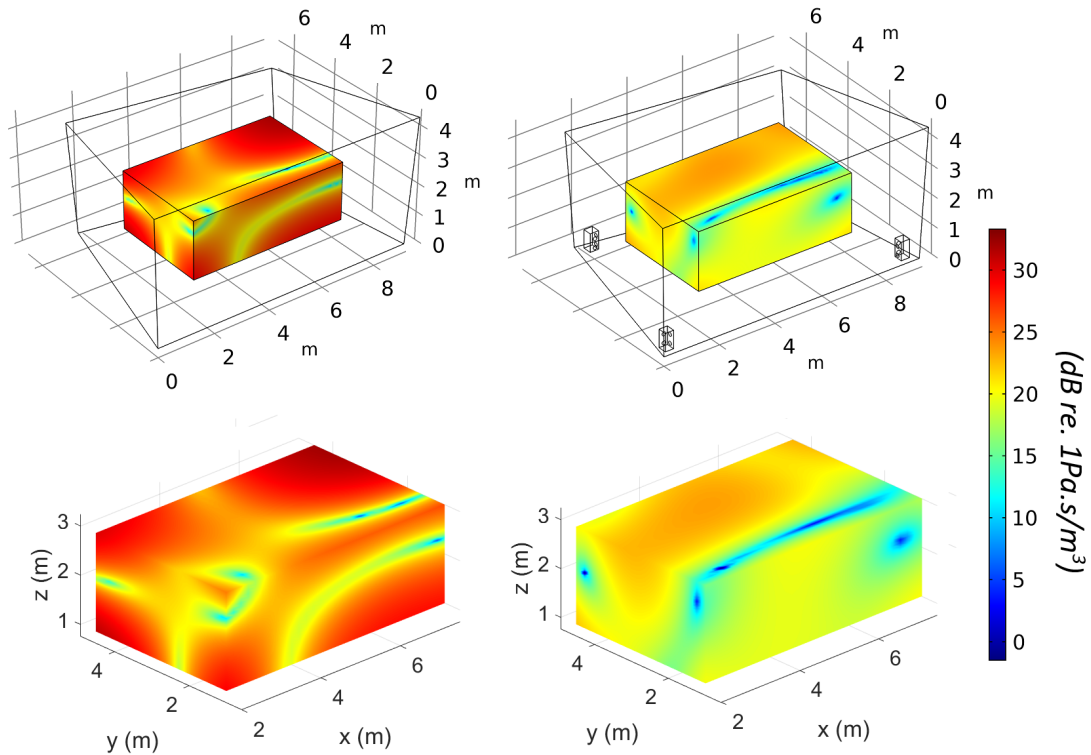


Figure 6.7 – The sound field before (left) and after (right) applying the electroacoustic absorbers are depicted precisely using the reconstruction algorithm (bottom) when compared to the simulation results (top) at $f=55.07$ Hz

captured the overall reduction in terms of the average sound pressure in the room when the absorbers are turned on. Moreover, similar to the room damping comparison in 5.1, the smoothing effect of the spatial pressure change between the nodes and anti-nodes of the mode shape is also rendered correctly.

In Fig. 6.8 we performed the same analysis in terms of the number of input microphones versus the accuracy of the reconstruction. It can be observed once again that the performance of the reconstruction can be improved as the number of microphones increases. The correlation value can be seen to improve both in terms of the average value but also the standard deviation of the reconstruction across the microphones in the 11x11x11 comparison grid in Fig.5.7.

From the analysis, it can be seen through simulation results that the sound field reconstruction framework shows promising prospects regarding its uses in evaluating the performance of active room absorption techniques. These spatial results can bring to light numerous aspects that some disperse individual measurements cannot provide.

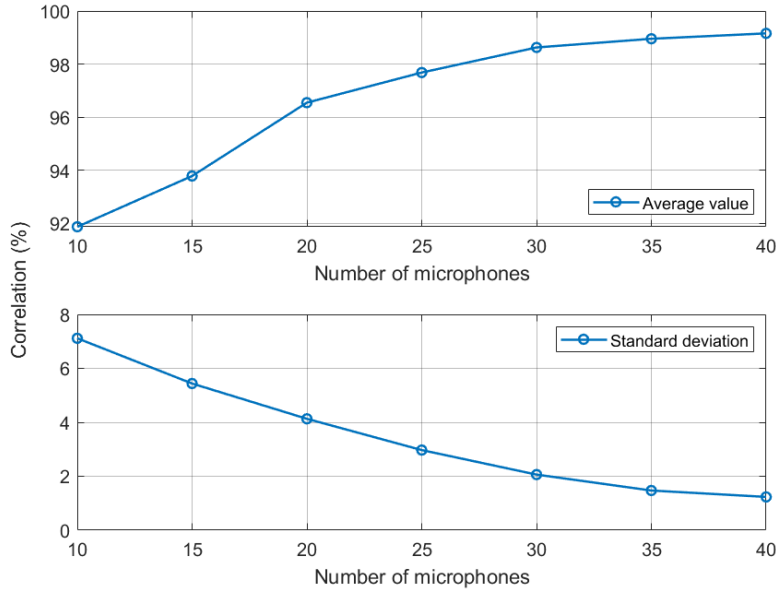


Figure 6.8 – Analysis of number of microphones vs correlation value

6.2.2 Experiment results

From the previous section, we have been able to observe the performance of the Electroacoustic absorber through a different lens where the changes to the spatial sound field can be visually rendered and analyzed. In this section, we aim to realize this numerical studies in real experiments.

The experiment, similar to Section 5.2, can be performed in the reverberation chamber at EPFL where the room design is the same as the room in the numerical model. Within the room, 4 identical Electroacoustic Absorbers are placed near the 4 corners of the room. The absorbers orientation is set to be exactly as the one in the simulation (facing towards the inside of the room, as in Fig. 6.9).

The reference source in the room is the same home made closed box subwoofer which can generate sound at low frequencies efficiently. In front of the source is the laser velocimeter that can measures the velocity of the diaphragm of the subwoofer. This velocity measurement can be used as the reference for processing the frequency responses of the room.

In this experiment, 35 measurements at different locations in the room are collected to be used both as the input and evaluation measurements. The measurements are done twice, once with the case where the absorbers are turned off and another when it turned on. The settings for the target impedance of the absorbers are set to be the 3-DOF model with the individual parameters exactly as listed in table 6.2. The microphone locations were tested in advance using the numerical model to make sure that they can captures well the modes of the



Figure 6.9 – Overall experiment set-up for sound field reconstruction with the Electroacoustic Absorbers.

room for sound field reconstruction.

For our first analysis, 25 microphones are used to reconstruct the sound field of the room and the other 10 are used for verification of the results. Figures 6.10 compares the reconstructed sound field at the very first mode of the room around 20 Hz using the SOMP method for eigenmodes identification. The reconstruction is once again performed on the rectangular assessment volume within the chamber. It can be seen that the reconstruction of the sound field and the mode shape behavior are highly in agreement with what we have from the numerical reference.

Once again, as explained in section 5.2, due to the difference between the numerical and practical case, we should compare these results in the visual sense and not the quantitative sense. It can be seen that there can be several differences when comparing between the actual reconstructed sound field and the numerically generated one. This discrepancy can be higher than in the experiments in Section 5.2 due to the fact that we now have the model of the Electroacoustic absorbers in the FEM study which can't be exactly tuned to match the ones in real life. Nevertheless, it can be seen that the changes in the sound field are correctly reproduced by the reconstruction framework. Both the level and the dynamic of the pressure in the room are reduced and depicted with high accuracy.

Another comparison was done at a higher order mode of the room using the reconstruction based on the supervised RFP global curve fitting method which can be seen in figure 6.11. In this comparison, the sound field was reconstructed at the 9th mode of the room which is at around 53 Hz.

Once again, it can be observed that the RFP method can also achieve good results even for such complex mode. The effects the absorbers have on the shape of the sound pressure especially

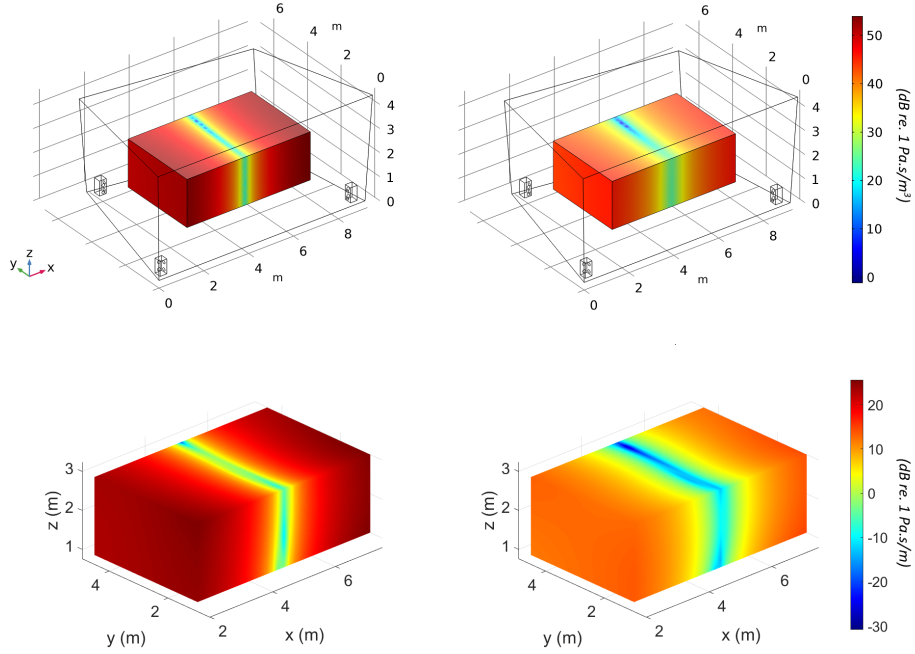


Figure 6.10 – The comparison of the reconstructed sound field using SOMP at the first mode (top) when the Electroacoustic Absorbers are turned off (left) and on (right) as compared to the equivalent situation simulated by the FEM model

at regions close to the nodal line of the mode are also similar to the numerical model. Similar to the previous case, there are also certain differences between the numerical study and real experiment. This shows the importance of the reconstruction framework as in real life it is impossible and impractical to 100 % replicate the situation in actual practice.

Using the other 10 microphones as the evaluation set, we evaluate the correlation value detailed in 5.2 for the frequency range of [15, 75] Hz and achieve the results listed in table 6.3.

6.3 Conclusion

Throughout this chapter, we have investigated the applicability of the reconstruction framework in analyzing the spatial performance of active absorber, in particular, the Electroacoustic Absorbers (EA) which is an impedance-based active control of loudspeakers in enclosures. The design of the EA was briefly introduced at the beginning in order to be able to replicate its theory in an FEM simulated model. The framework in Chapter 4 was then used to reconstruct the sound field of the room using the same strategies mentioned in Chapter 5 but this time with the presence of 4 EAs with each at a corner in the room. The reconstruction results showed promising recovery results with the Pearson Correlation percentage remains to be very high and the visualization of the modal behaviors of the room can be seen to captured

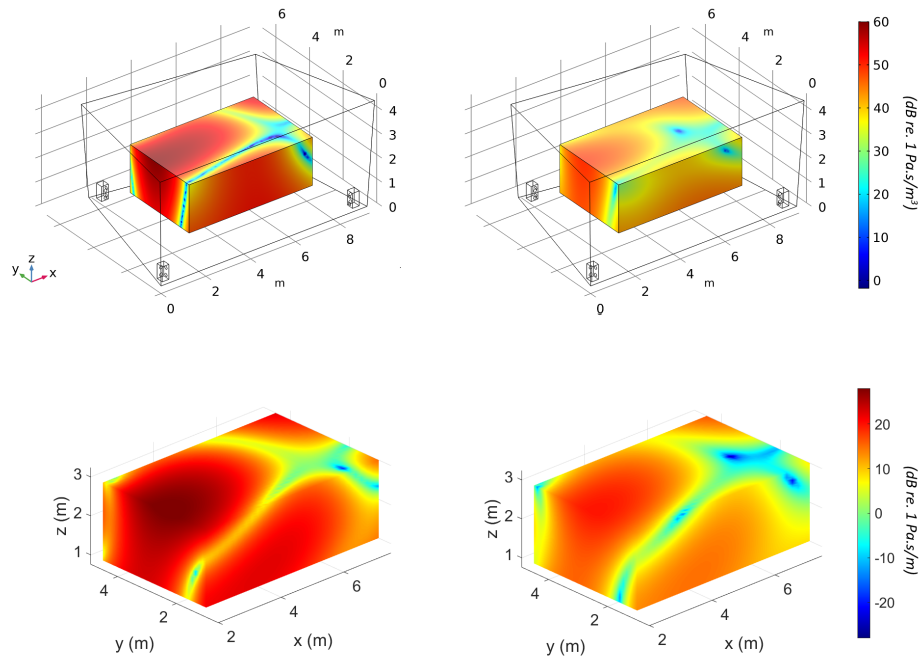


Figure 6.11 – The comparison of the reconstructed sound field using RFP at the 9th mode of the room (top) when the Electroacoustic Absorbers are turned off (left) and on (right) as compared to the equivalent situation simulated by the FEM model

fully its changes when compared with the case of an empty room.

Proceeding with the same procedure, the reconstruction process was then performed for the real reverberation chamber at EPFL with the same placements of the EAs at the corners. Once again the reconstruction framework returns highly accurate recovery of the sound field in the room where the changes in the spatial modal behavior can be clear observed in all cases of testing. Overall, this shows that the reconstruction framework that we developed can be effectively used not just for an empty room but also when the room is occupied by not just passive but active control elements such as the Electroacoustic Absorber. This opens door to future application of the framework to assess the spatial performance of multiple types of active absorbers in the room. Furthermore, it could also be potentially used to compare the spatial performances between different types of active absorbers.

In Chapter 7, we will investigate an even more intensive application of the reconstruction framework where it can be used to effectively tune the control parameters of individual resonator in the Electroacoustic Absorbers control design.

Table 6.3 – Correlation value evaluation of the reconstruction from real measurements for different cases and methods

| Case | Method | Average | Standard deviation |
|------|--------|---------|--------------------|
| OFF | SOMP | 97.4 % | 1.74 % |
| | RFP | 97.1 % | 1.91 % |
| ON | SOMP | 97.0 % | 1.92 % |
| | RFP | 96.7 % | 2.15 % |

7 Sound field reconstruction application in room modes equalization

In the previous chapter, we have been able to see the application of the sound field reconstruction framework in analyzing the spatial performance of active absorbers in a non-rectangular room. The reconstruction has been shown to be reliable even in the cases where there are multiple absorbers in the room with varying individual responses. By doing so, not just that the performance of the absorber can be analyzed based on familiar metrics such as the modal decay time but furthermore, the equalization effects it has on the spatial modal behavior of the room can also be graphically presented.

In this chapter, we would like to investigate an even more advanced use of the reconstruction framework in a sense that its reconstruction results could assist the design process of the electroacoustic absorber. In order to do so, in Section 7.1 we first look for possible evaluations or metrics that can be used to analyze the equalization aspect of a frequency response curve which could be useful for the design of the target impedance for the Electroacoustic Absorber. In order to accomplish this, in Section 7.1.1, we first review the theory behind a target-impedance-based control and study its effects in the same reverberation chamber mentioned in previous chapters. We also briefly review how the target impedance has been computed from numerical data in past attempts to damp room modes. From there, Section 7.1.3 experiments with new promising metrics that can be acquired in practice thanks to the sound field reconstruction framework. After testing these metrics with numerical data to confirm their validity and efficacy, Section 7.2 applies these metrics in combination with the sound field reconstruction framework to propose a new way to calibrate and locate the optimal control parameters for the 1-DOF resonator model of the Electroacoustic Absorber. This could potentially open doors to promising new directions in the field of room modes equalization in the future.

7.1 Room modes Equalization metric based on Sound Field Reconstruction

We first introduced the subject of the Electroacoustic Absorbers in Chapter 6 without explaining much about its founding theories. In this section, we analyze in more details the subject of target impedance and furthermore briefly review the reason behind the choices of the control parameters for the target impedance of the aforementioned electroacoustic absorbers (table 6.2) which has been researched in [101]. From this analysis, we will discuss how the existence of a robust sound field reconstruction framework may assist the control design process for the electroacoustic absorbers in particular and future prospective active absorbers.

7.1.1 Optimal specific acoustic resistance

1-D analysis

In [60], a 1-D analysis of an FEM model of a duct with a square cross-section was performed to study the dependent of the target impedance with respect to the effective area of absorption. The duct has one perfectly rigid end, and the other end is a passive absorber with the area equal to the rectangular cross section of the duct. Appointing a value of $Z_{s_{abs}}$ for the specific acoustic impedance of the absorber, it has been found that for the value of $Z_{s_{abs}} = \rho c$, the modal decay time of the longitudinal mode of the duct approaches 0 due to the fact that the duct approaches that of an open terminal condition on the side of the absorber.

The situation changes, however, for the cases when the effective absorbing area of the absorber (S_{abs}) is smaller than the actual cross section of the rectangular duct (S_{duct}). It was then argued that for each value of S_{abs} (granted that S_{abs} is not too small), there exists an optimal value that can results in maximum modal damping.

In Fig. 7.2, the evaluation of the modal absorption performance of the circular absorber was analyzed for the first and third longitudinal mode of the duct using the modal decay time, which has been defined previously as:

$$MT_{60n} = \frac{3 \ln(10)}{\delta_n} \quad (7.1)$$

in which δ is the imaginary part of the eigenfrequency ($\omega_n + j\delta_n$).

From the figure, we can observe that depending on the value of the area of the absorber, there exist a value of the impedance of the absorber such that the modal decay time is minimized (under the circumstances and boundaries of the study). These values lie below the value of the specific acoustic impedance of the medium $Z_{s_{air}} = \rho c$. Moreover, it can also be seen that this so-called optimal value is generally different for the two modes under study and this behavior

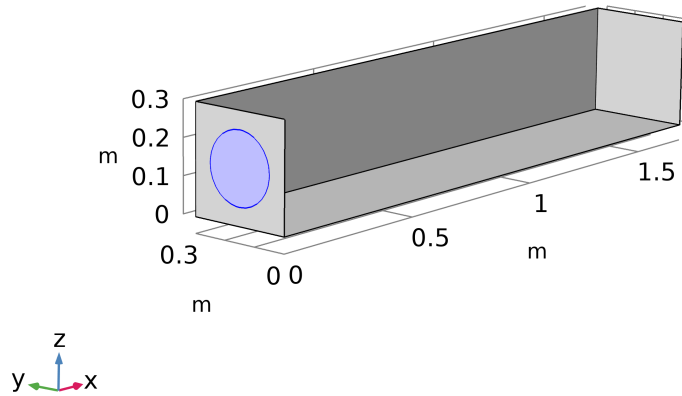


Figure 7.1 – The model of the duct under study ($S_{duct} = 900\text{cm}^2$, $L_{duct} = 1.7\text{m}$) and the circular absorber with a variable radius and specific acoustic impedance at one end of the duct.

has been proven to extent to higher-order modes as well.

Analysis in a 3-D enclosed room

The study of the optimal impedance for the 1-D case lays the foundation for the extension of the scope to the 3-D case in a room, with the hope that it will provide important clues on how to implement the target impedance for the electroacoustic absorbers. In [101], the study was further performed for a medium sized rectangular room.

In this section, however, we will replicate this study in the FEM model of the reverberation chamber, using the electroacoustic absorber model with the same dimension as the existing ones in EPFL previously presented in Fig 6.2. This will allow us to analyze multiple similar aspects using the particular room that we have been studied since the previous chapters. Four of these electroacoustic absorbers are placed in each of the 4 corners of the room (Fig.7.3).

In this study, we perform the eigenfrequencies study in COMSOL while varying both the radius of the membrane of the 16 enclosed loudspeakers as well as their specific acoustic impedance (which is kept purely resistive). The radius of the speaker is swept between 7 cm and 12 cm (which is a considerable limit considering the size of the actual enclosure). The impedance of the loudspeaker surface is set to vary between $0.01\rho c$ to $0.2\rho c$ which is likely to be where we can observe the optimal region of the MDT. One aspect to notice is that in this case we purely experiments on the so-called passive model of the electroacoustic absorber where there were

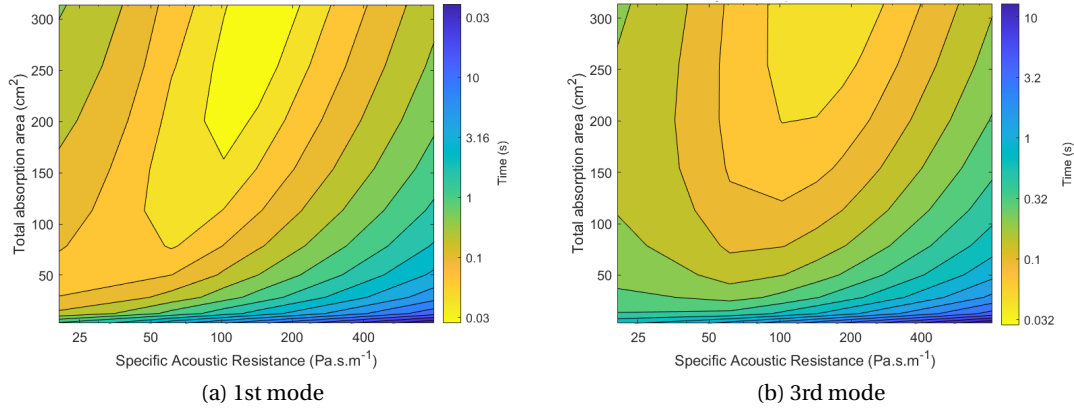


Figure 7.2 – Modal decay time analysis between different combination of the effective absorption area and the purely resistive impedance of the absorber at one end of the rectangular duct. Total absorption area is equal to the total physical area of loudspeaker elements.

no design of the mechanical elements of the speaker in COMSOL. The impedance of the are purely resistive and assumed to be constant across all frequencies.

From the eigenfrequencies, similar to the 1-D case, we gather a list of eigenfrequencies of the room which includes both the modal angular frequencies ω_n and the modal damping factor δ_n . We then analyze the behaviour of the Modal decay times of different modes of the room with respect to the corresponding total absorbing area and impedance of the loudspeakers. Figure 7.4 presents the modal decay time analysis for 6 different modes of the room and their respective relationship with the total absorbing area and specific acoustic resistance of the loudspeakers.

As the room has a non-rectangular shape, most of these modes can't be easily identified as an axial, tangential or oblique mode. The analysis in Fig. 7.4 presented respectively the results for the 1st mode (≈ 20 Hz), 2nd mode (≈ 27 Hz), 3rd mode (≈ 35 Hz), 6th mode (≈ 45 Hz), 12th mode (≈ 59 Hz) and 14th mode (≈ 66 Hz).

We can observe from the figure that the trends on all six modes of the room regarding the modal decay time are similar. Typically, there exists an optimal region (for the ranges within our studies) for the absorption area and the specific acoustic resistance of the absorbers to acquire a favorable modal decay time. In more details, if the specific acoustic resistance lays on the right side (higher than) with respect to this optimal region, increasing the total area of the membrane tends to improve the modal decay time of the room. Interestingly enough, within the ranges of our study, when the resistance lies relatively on the left side of this domain, increasing the area of the absorbers offer barely any changes to the performance not to say that it might even increase the modal decay time.

A better way to phrase the relationship between the total effective absorbing area and the

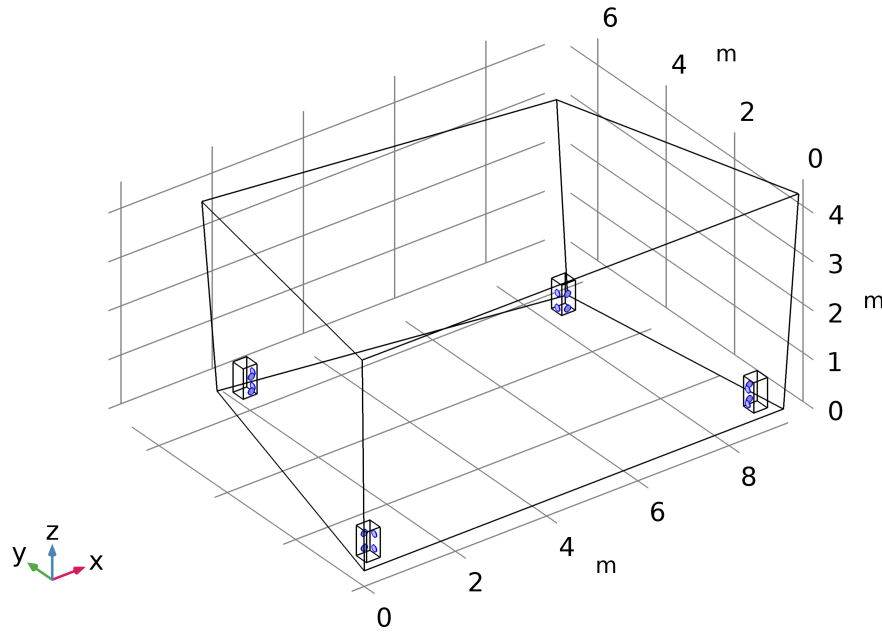


Figure 7.3 – The model of the eigenfrequency study for the reverberation chamber for variables parameters, namely, the radius of the 16 highlighted speakers of the four Electroacoustic absorbers and their specific acoustics impedance.

respective specific acoustic impedance of the absorber within our region of study is that for a specific value of the area, there exists a region of impedance within which the modal decay time can be optimally reduced. Similarly, for a specific value of the resistance, there also exists a certain range of area such that the modal decay time can be optimally reduced.

Another characteristic that can be observed from this analysis is that although they share the same behavior with regards to these two factor, this so-called optimal regions are located at different combination of the area and resistance for different modes. For example, analyzing the absorption performance for $S_{abs} = 2500 \text{ cm}^2$, it can be seen that the optimal resistance for the 1st mode lies around $14 \text{ Pa} \cdot \text{s} \cdot \text{m}^{-1} (\approx \rho c / 29)$ while this value for the 2nd mode is at around $20 \text{ Pa} \cdot \text{s} \cdot \text{m}^{-1} (\approx \rho c / 20)$. For the 12th mode, this value is located even further away with $R_{s_{abs}} = 37 \text{ Pa} \cdot \text{s} \cdot \text{m}^{-1} (\approx \rho c / 11)$. This indicates that assuming we have a fixed area of absorption, even though for each mode there exist a limited range of optimal value for the specific acoustic resistance, this optimal value cannot be generalized for all the modes of the room.

Another observation from this investigation is that for a fixed area of absorption, the effect that the optimal value of specific acoustic resistance has on the modal decay time also differs between different modes of the room. For instance, when $S_{abs} = 2500 \text{ cm}^2$, the best modal decay time that can be achieved at the 1st mode is 1.17 second, while this value for the 14th mode is roughly 2 seconds. This value of modal decay time generally increased as we go to

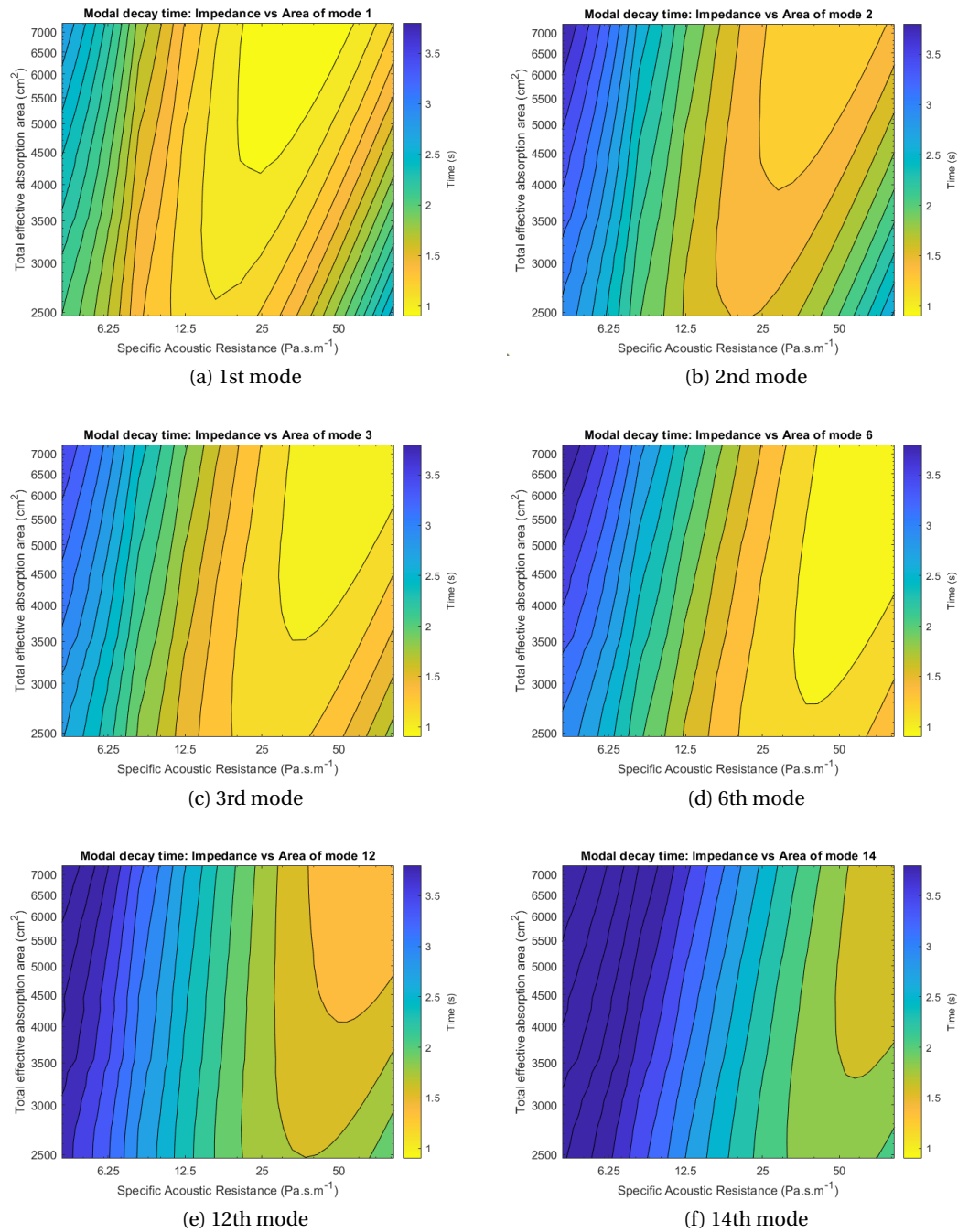


Figure 7.4 – Color mapping of the modal decay time of several modes in the reverberation chamber with respect to the specific acoustic impedance and the total effective absorption area of all four absorbers placed at four corners of the room.

the higher order mode but there also exists exception. For example, from another evaluation which is not listed in the figure, the optimal modal decay time for the 5th mode is at 1.2 second while for the 8th mode it is at 1 second. It is expected that there will be even more instances like this when the design of the room is more complex or when there are other objects in the room with different absorption coefficients that are frequency-wise diverse.

From these analysis, we can see that for a fixed value of S_{abs} there does not exist a single optimal value of $R_{s_{abs}}$ that can absolutely minimize the modal decay times for all the modes of the room. Furthermore, different room modes have different way of reacting to the impedance that can be applied at the membrane of the loudspeakers. This diversity of behavior are, furthermore, difficult to predict as it depends on numerous deciding factors. It is also worth noting that in real life, it is not practical to design a target impedance that is purely resistive for a loudspeaker. Hence, studies on the effects of the imaginary part of the impedance to the overall results require further investigation in the future.

7.1.2 Past optimization attempts

In the previous section, the subject of optimal resistance has been introduced and analyzed for the 1-D and 3-D cases. One important conclusion that can be made is that for the 3-D case in a room, it is not straightforward to find a single value of the impedance that can optimally bring the modal decay times of every mode in the room to a minimum. This means that several compromises as well as optimization should be taken into account for a target-impedance-based design.

In this section, we briefly review the optimization process for the Electroacoustic absorber and what had previously been done prior to our research to compensate for the aforementioned obstacle regarding the optimal impedance. At the end of the analysis, it shall become more evident on how a sound field reconstruction framework might be able to provide a different angle on the control design process of the electroacoustic absorbers in particular and active room mode equalization process in general.

Due to the varying so-called optimal specific acoustic impedance of different modes of the room, the research in [99] proposed an optimization regime that attempts to come up with a target impedance that directs towards a compromise in a sense that, though not every modes can be brought to its theoretical optimal modal decay time, the target impedance nevertheless can be formed into a shape in the frequency domain that compromises this criteria for a large number of modes within a frequency range of interest.

In order find this target impedance curve for a fixed total area of absorbers (which is equal to the total areas of 16 loudspeakers with S_d as listed in table 6.1), first, the optimal resistance for each and every modes of the room can be superimposed on the same plot. In order to generalize this absorber design for multiple rooms, the aforementioned research took into consideration 3 different model of a small and two medium sized rectangular rooms. Moreover,

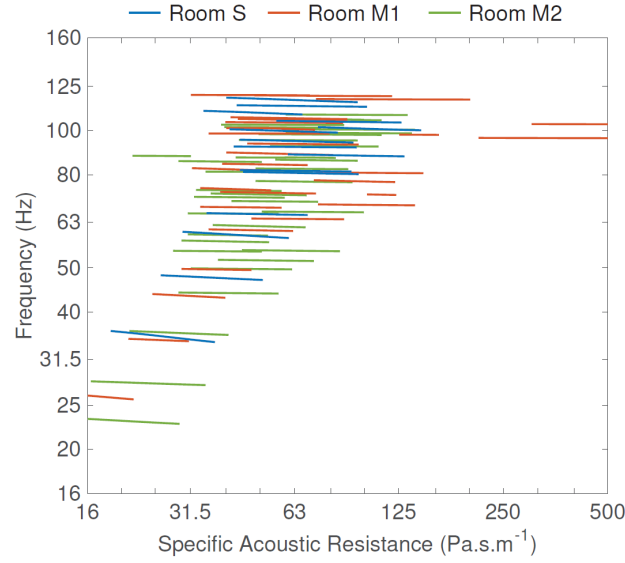


Figure 7.5 – Plot produced by E. Rivet consisting of optimal ranges for the value of specific acoustic impedance for multiple modes of three different room models. The ranges are plotted with respect to the modal frequencies, so as to later form a bounded optimal region for the target impedance curve in the frequency domain.

to leave some room for tolerance, each optimal resistance is expressed not in terms of only a single value, but a small varying range which allows the modal decay times to fluctuate within a 66 ms boundaries. Figure 7.5, which was produced in [99], plotted this large set of optimal ranges for the specific acoustic resistance. The objective was then to create a target impedance curve that stays as close as possible to all these values of resistance across the frequency range of interest.

An optimal objective function was then created to best matched the impedance of the 1-DOF resonator model of the control to this target impedance regions. This objective function will help acquire the control parameters for the 1-D electroacoustic absorber design (which includes ν_1 , ν_2 and R_{st} , see Section 6.1.2). These parameters can then be used to deduce the parameters for the 3-DOF absorber model that has been listed in table 6.2 which is the same control that has been used for the existing Electroacoustic absorbers at EPFL. A multiple degree of freedom model, as explained in [99], will help the target impedance curve to possess a wider bandwidth as well as keeping its imaginary part as close to zero as possible on a wider bandwidth. This has been proved to be beneficial in [37] for a targeted resistance model. Figure 7.6 shows the improvement that can be achieved when using a multiple degree of freedom model.

This 3-DOF model for the target impedance has been shown to be able to reduce the modal decay time of a room by a considerable amount in a wide bandwidth as can be seen from a few examples in the real sound field reconstruction results in Section 6.2.2. However, from a prac-

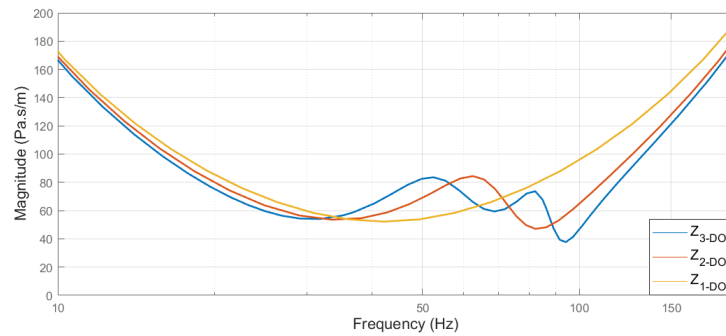


Figure 7.6 – The 1, 2 and 3 degree-of-freedom settings of the impedance which share the same asymptotes but a wider bandwidth can be achieved using the higher degree-of-freedom models.

tical point of view, several comments could be made regarding the design and optimization process.

The first aspect that should be pointed out is that the entire set of data of modal decay times that was used for the optimization process purely comes from simulation. Although the data has been simulated from 3 different rooms to increase the robustness, at the end of the day, it is extremely difficult to characterize a real room and its acoustic properties. Due to this reason, in the real case, not just the optimal values of acoustic resistance in Figure 7.5 might change, but also the bounds for each of their tolerance ranges could also be altered.

Another comment that can be made about the optimization process is that it does not take into account the imaginary part of the impedance when the sweeping parameters study was done. The reason is that it would increase the difficulties of the analysis as there will now be other sweeping parameters such as the center frequencies and bandwidth. Combining this with the plotting and manually sorting the modal decay times in three different room design before applying the optimization would make this a time consuming and inefficient process.

Finally, the modal decay time is a very useful parameter to assess the damping of a mode in a room. However, as has been seen in the previous analysis, when grouping various modal decay times at different modes altogether, it becomes more difficult to see the big picture as there isn't a global evaluation metric to combine them together. This becomes even more difficult in cases where a comparison between different active room treatment methods is needed to be done. For instance, if another 3-DOF target impedance was designed instead of the one in Table 6.2, how should we compare them in terms of performance on a wide bandwidth and not just for a single mode?

7.1.3 Room modes equalization metric using sound field reconstruction

All the aspects that had been raised at the end of the previous section become even more intriguing now that we have an available sound field reconstruction as a tool of assessment of the room spatial sound field at each and every frequencies under our scope of interest. Parts of these aspects are specifically for the particular control scheme of the Electroacoustic Absorber, but there are also common issues that can be generalized for other room modes treatment methods as well.

In this section, we study towards new metrics that can be extracted from the sound field reconstruction framework which might be useful in assessing the room modes equalization performance of room treatment methods (both active and passive).

To begin, we address the idea of room modes equalization and what it actually means. In an extreme situation, room modes equalization would mean that the effects of room modes in the room no longer exists and the effect of the room on the listening experience becomes zero. One of the way to express this is that the room frequency responses in the room becomes a flat line that does not favor any particular frequency in the room. This would also mean that the spatial responses of the room are, in a way, neutral, where every point in the room has the same reaction to the sound created by the source in the room.

When analyzing the flatness of the response curve in the frequency domain, one of the metric that could be mentioned is the standard deviation of a curve which indicates how much the curve deviates from its average value. In the frequency domain, it means how much the frequency response deviates from its average value, calculates by averaging all the responses at every discrete frequencies.

In order to test the viability of this idea, we perform a modified version of the sweeping parameters study from Section 7.1.1. Similar to the previous study, we vary the area of the 16 membranes in this case along with its specific acoustic resistance. Knowing the ranges of the so-called optimal resistance for each mode beforehand, we sweep the resistance within the range of $\rho c/50$ and $\rho c/4$ and the radius of the speakers between 7 cm and 12 cm. The biggest difference in our numerical analysis is that our output data is much larger this time. The current study instead of collecting the eigenfrequencies of the room, will gather the frequency responses of a 6x6x6 3-D grid of measurement points which sample the inner rectangular area that we have mentioned in Figure 5.5. The measurement grid can be seen in Figure 7.7.

With the response data for the 216 measurement points in the numerical study, we now calculate the standard deviation value for each and every points. For each of the combination of S_{abs} and $R_{s_{abs}}$, we end up with 216 values of standard deviation σ_i with $(i \in [1, 216])$. Summing the results up and averaging them for each of the combination will end up with an average standard deviation value of $\overline{\sigma}_{(S_{abs}, R_{s_{abs}})}$.

Plotting the standard deviation with respect to the two varying parameters will give us Figure

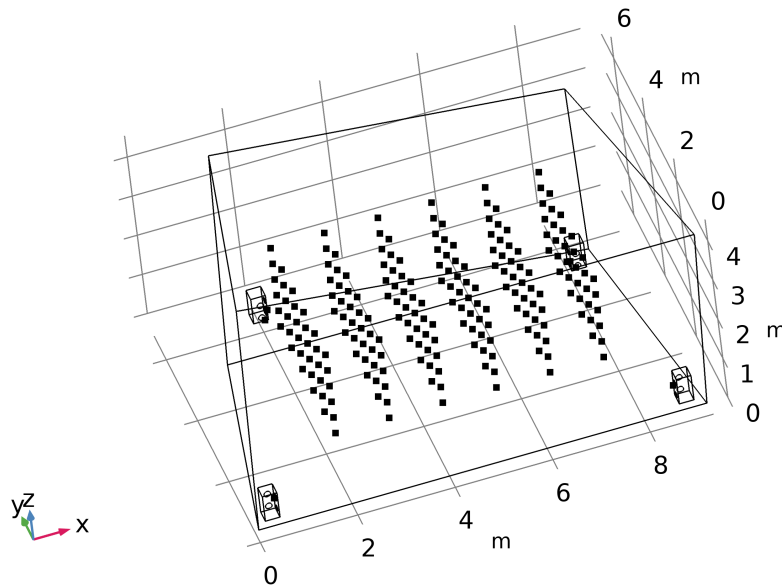


Figure 7.7 – The 216 points of measurement grid used for the sweeping parameters study in the numerical simulation.

7.8.

As it can be seen, there is no discernible pattern in the map of the average standard deviation and there is not much coherent with what we know about the modes in the room. In fact, the map does not show a pattern that can represent the equalization of the room at all because from Section 7.1.1 we know that for the reverberation chamber FEM model, the optimal values for the resistance for every mode are all higher than $10 \text{ Pa} \cdot \text{s} \cdot \text{m}^{-1}$. This can become even more clear if we observe Fig. 7.9 when we compare two different responses at the same point. One has a higher standard deviation than the other. However, the one with the lower average standard deviation is actually the one with a sharper rise and drop of pressure than the other.

In order to eliminate the bias, we take into account several modification to the evaluation. The first modification is to take into account the average of the responses instead of just the standard deviation alone. This could potentially reduce the bias coming from the difference in terms of energy between different settings.

Another possible modification that can be made into the calculation is to take into account the fact that calculating the average of the response for a very wide frequency range is not a good idea, considering the fact that there is a large number of modes within the range. Moreover, the absorption characteristic of the room frequency-wise might also affect the baseline of the responses frequency-wise. This is why we opted for the moving averages and moving standard deviation instead of just a single universal value for the entire range.

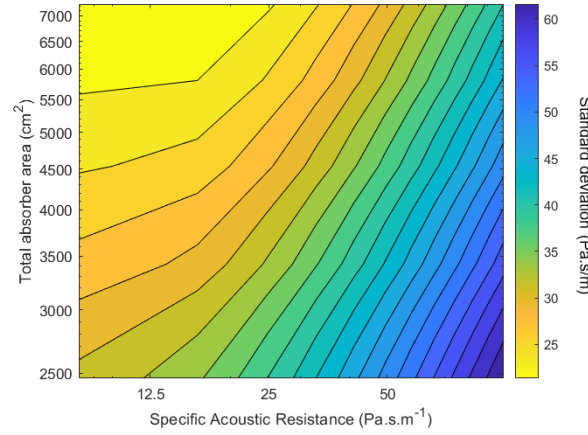


Figure 7.8 – Average standard deviation of the Room frequency responses at the 6x6x6 sample points with respect to the total effective absorption area and the specific acoustic resistance at the 16 speaker membrane

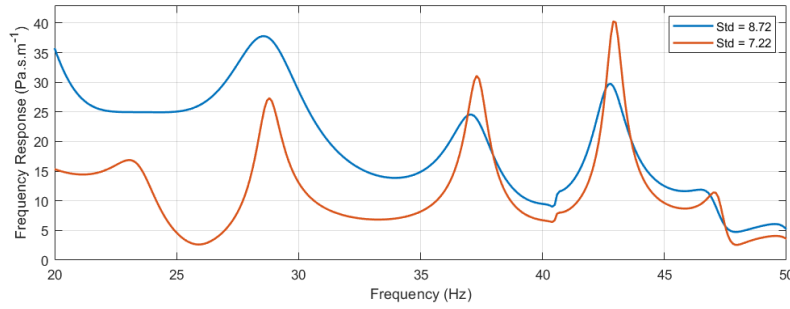


Figure 7.9 – A comparison between two different responses with different average standard deviation values which shows that the standard deviation could be a misleading metrics in analyzing the equalization aspect of room response.

From these remarks, two different methods of evaluating the RFR curve are proposed. The first method calculate the ratio between the moving standard deviation and moving average. The second method instead focuses on evaluating the so called Root mean square error of the response when compare to its moving averages.

For the first method, we first impose a modified version of the moving average (μ_{mod}) and standard deviation σ_{mod} where the length of the moving window changes along the frequency axis. The reason behind this is that in a room, the number of modes increases rapidly with frequencies and hence using a fixed size window would not be appropriate as the number of modes included in a window would increase drastically. This is why we opted for a window

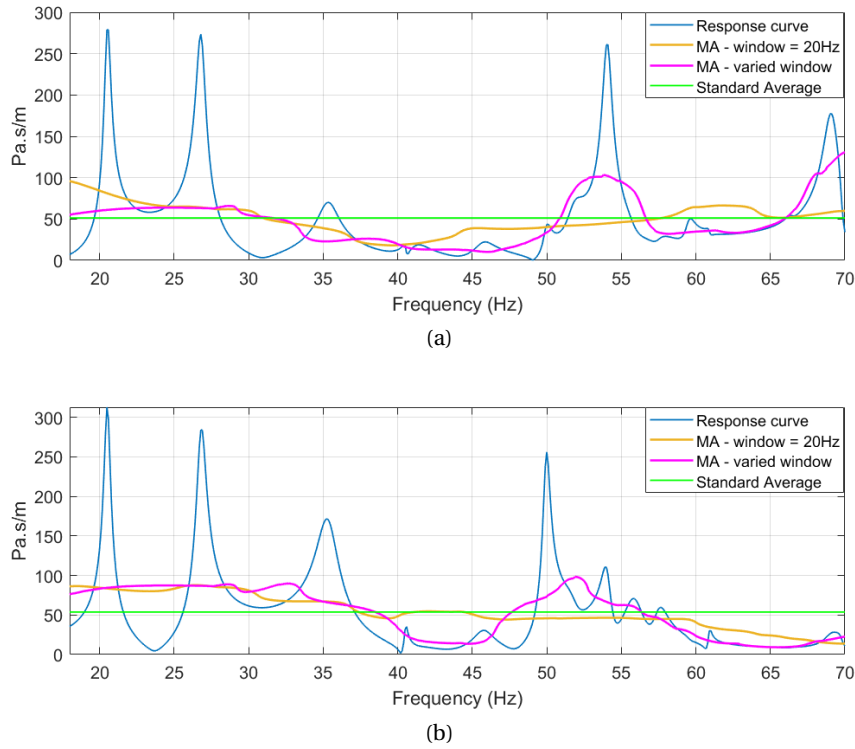


Figure 7.10 – Comparison of different averaging methods with respect to the original response curve.

with changing length using the modal density estimation for room acoustics:

$$\frac{dN}{df} \approx \frac{4\pi V}{c_0^3} f^2 \quad (7.2)$$

in which V is the volume of the room and c_0 is the sound celerity. Using the inverse of this value will give us an estimation on the distance between two consecutive modes at certain frequency f . Using this value as the variable length of the window, we calculate the average of the value σ_{mod}/μ_{mod} across the frequency axis for each of the evaluation point and then averages this value across all evaluation points in the $6 \times 6 \times 6$ grid to get a single representative value for each combination between S_{abs} and R_{sabs} .

The second method takes a slightly different approach in that it uses the modified moving average mentioned above as the baseline for the Frequency response. Figure 7.10 plotted 2 examples of the modified moving average compared to a standard constant average as well as a moving average with fixed window length of 20 Hz with reference to the original response curve. From this figure, it becomes clearer why a constant average over a wide frequency would not be a good idea as there are multiple modes in the frequency range. The moving

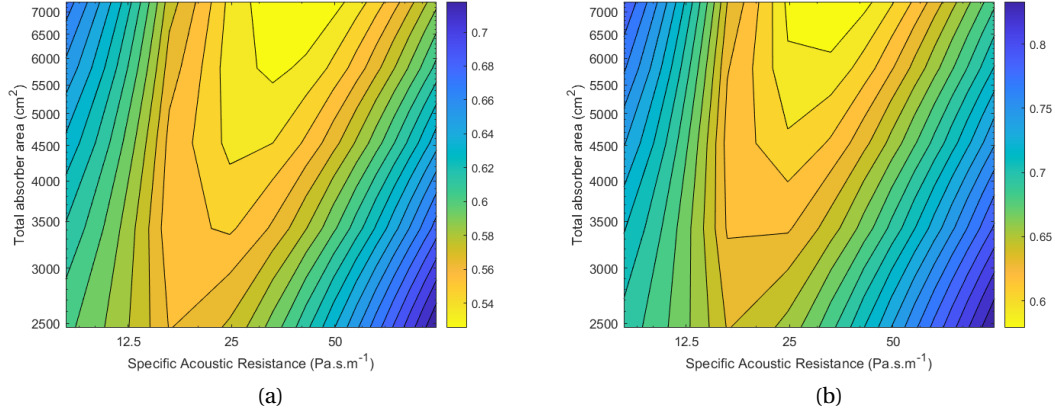


Figure 7.11 – Evaluating the sweep parameters study using two different metrics: a) Average relative moving standard deviation b) Average Root Mean Square Deviation (RMSD).

average with fixed window predictably does a better job than the constant average, however, frequently, due to the unbalanced response curve, the moving average find itself higher than some of the peaks of the modes which is not ideal. Finally, out of the 3 variations, it should be noticed that the moving average with a varied window length possesses the tendency to stay below most of the peaks of the room modes. It can be seen that this function can be effectively used as the baseline of the RFR. The response can then be evaluated by calculating its deviation from this baseline reference. The metric we used in this case is similar to the Root mean square error (RMSE) metric which has been used in common estimation problem. In this case, the other name of it, which is the Root Mean Square Deviation (RMSD) is more fitted to be used in our case:

$$\text{RMSD} = \sqrt{\frac{\sum_{f=f_{\min}}^{f_{\max}} (S(f) - \mu_{\text{mod}}(f))^2}{L_f}} \quad (7.3)$$

The relative RMSD can then be found by dividing the RMSD by the modified moving averages $\mu_{\text{mod}}(f)$. Similar to the previous method, averaging this value across the frequency and across the 6x6x6 evaluation will give us a single value to represent each combination of S_{abs} and R_{sabs} .

Figure 7.11 details the performance evaluation of our sweep parameters study in FEM using the two methods mentioned above.

The first observation that can be made is that the trends from the two different methods of evaluation are quite similar to each other. Moreover, they exhibit high resemblance of the modal decay times plot for a single mode in Figure 7.4. The first thing to remember is that the

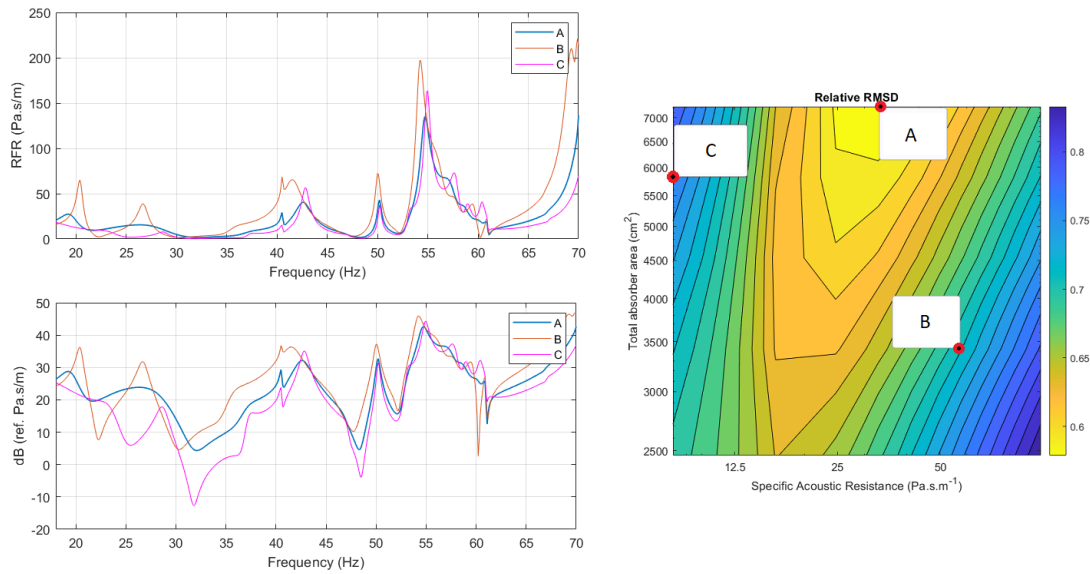


Figure 7.12 – Frequency response (both absolute value and dB value) at a single point in the room under different combination of the sweep parameter study as compared to their location on the color mapping of equalization.

evaluation we are investigating, unlike the modal decay time plots, concerns not just a single mode but a wide frequency band. This means that the plots in Figure 7.11 will exhibit values that has taken into account the compromise between all the modes within this frequency band, as it has been shown in Section 7.1.1 that different modes in the room will have different regions of optimal resistance. This becomes even clearer when we analyze the values of the optimal resistance in this case. For a fixed value of the total absorption area, the optimal region for $R_{s_{abs}}$ always lies somewhere in between the optimal values for each and every modes.

Another observation is that the relations between the absorption area and the specific acoustic resistance are also maintained across the two evaluation and exhibit the same behavior as when we analyze for a single mode. In order to verify if the two plots make sense and do point us to an agreeable level of satisfactory, Figure 7.12 plots the responses at the same position for three different combination between S_{abs} and $R_{s_{abs}}$ and their respective locations on the sweep study plot. It can be seen that the response for case A is the most equalized out of the three curves which reflects the location of its metric on the RMSD map. The strength of this evaluation is that it takes into account a large amount of responses within an area of interest, however, even for some specific individual comparison like the one in the figure, we can still see that the metrics we used perform quite well in characterizing these combination on the equalization aspect. Moreover, now that we have developed a robust sound field reconstruction algorithm, this evaluation can be applied to the reconstructed sound field from a set of real measurements in the room.

The two metrics that we used for the evaluation in Figure 7.11 are just two of the multiple

possible ways to evaluate the equalization status response curves. Other criteria such as energy balance between bandwidth, average deviation in space, average total dynamic range within a bandwidth... could also be investigated for use in the future depending on specific objectives. From Figure 7.11, certain small differences between the two metrics could be noticed for example, the RMSD criteria seems to have a finer quantization for the combination near the optimum region compare to the relative moving standard deviation which is a little more relaxed. Nevertheless, these two metrics do seem to exhibit very similar categorization of the sweep study results.

These equalization evaluations could potentially be used in comparing the performance of different room equalization methods within a bandwidth of interest, which is something individual MDT comparison would find it difficult to do. Another advantage of this space-wise evaluation is that it could also be done for a specific region in the room. For example, in music listening rooms or small music chambers, it could be performed specifically for the audience targeted area.

Another use for metrics like these is to give suggestion as well as indication for acoustic researchers when designing active sound absorbers in the room. The active control of such active absorbers could then be tested and fine tuned based on these evaluation. For instance, when we consider the design of the electroacoustic absorber in Section 6.1.2, this global evaluation could also be potentially used to fine tune the properties of its individual resonator such as resonant frequency, bandwidth as well as the specific acoustic resistance. The next section will focus on experimenting with this idea.

7.2 Target Impedance analysis and fine tuning using Sound Field Reconstruction

In the previous section, we have introduced two different metrics that might prove helpful in analyzing the equalization aspect of the frequency responses of the room. In this section, we focus on experimenting this very idea on the specific case of the Electroacoustic Absorber control design.

7.2.1 Prototype

To help target the low frequencies of the first few modes of the room, we built a new prototype of the electroacoustic absorbers which have a few differences when compared to the design in Figure 6.2. Built into the new prototype are 8 instead of 4 loudspeakers to make use of all its surface area. They are divided into 2 levels with 4 speakers on each level. Instead of dividing the enclosure on each level so that each speaker has its own enclosure, we also experiment with the idea of allowing every 4 of them to share the same enclosure so that the volume of the enclosure can be more than doubled without increasing the overall size of the wooden box. The overall size of the prototype is $0.31\text{m} \times 0.31\text{m} \times 0.64\text{m}$. The loudspeakers we used

Table 7.1 – Small signal parameters of the closed-box loudspeaker system for the new prototype of the Electroacoustic Absorber

| Parameter | Notation | Value | Unit |
|-----------------------------|----------|--------|-----------------------------------------------|
| Enclosure volume | V_b | 20 | dm^3 |
| Moving mass | M_{ms} | 14.67 | g |
| Mechanical resistance | R_{ms} | 1.31 | $\text{N} \cdot \text{s} \cdot \text{m}^{-1}$ |
| Total mechanical compliance | C_{mc} | 397.34 | $\mu\text{m} \cdot \text{N}^{-1}$ |
| Effective piston area | S_d | 151 | cm^2 |
| Force factor | Bl | 5.98 | $\text{N} \cdot \text{A}^{-1}$ |
| Air density | ρ | 1.2 | $\text{kg} \cdot \text{m}^{-3}$ |
| Air celerity | c | 343.86 | $\text{m} \cdot \text{s}^{-1}$ |

for the prototype are the Peerless SDS-P830657. The absorber was built with a removable top panel using simple locking mechanism on the 4 sides to allow physical as well as electrical modifications in the future. The overall design of the absorber can be seen in Figure 7.13. The parameters of the new prototype can be found in Table 7.1.

A voltage controlled current source was used to drive the loudspeakers. The active control of the driven current of the loudspeakers was done using Speedgoat Real-time Baseline target machine which includes IO397 I/O Module. The incorporated control scheme can be built directly in Simulink Real time. The transfer function of the controller Θ from Figure 6.1 can then be modified in real-time to adapt to different control model.

Two identical prototype of the Electroacoustic Absorbers in 7.13 are built and can be put in the reverberation chamber. To briefly test its performance, the 3-DOF presets in Table 6.2 are used to feed into the controller. The performance are quite effective and similar to what can be achieved from the previous prototype.

7.2.2 Optimal resistance in real experiment

Before using the reconstruction framework to calibrate the control of the absorbers, we first perform a small series of experiments on target resistance to see if it really does exist for room modes. This not only can help us assess the FEM studies of target impedance, but also can give us some ideas on the bounds of the impedance value for different room modes.

As mentioned in Section 7.1.1, a purely resistive target impedance is not practical to apply to the control of active absorbers. In fact, in the case of a single degree of freedom resonator, for certain values of the resistance R_{st} in Eq. 6.12, depending on the physical design of the absorber as well as the control system, there exist an upper bound for the bandwidth of the realized impedance before the control gets itself into instability. Usually, in practice, as R_{st} gets lower, the bandwidth would also have to be narrowed to avoid instability.

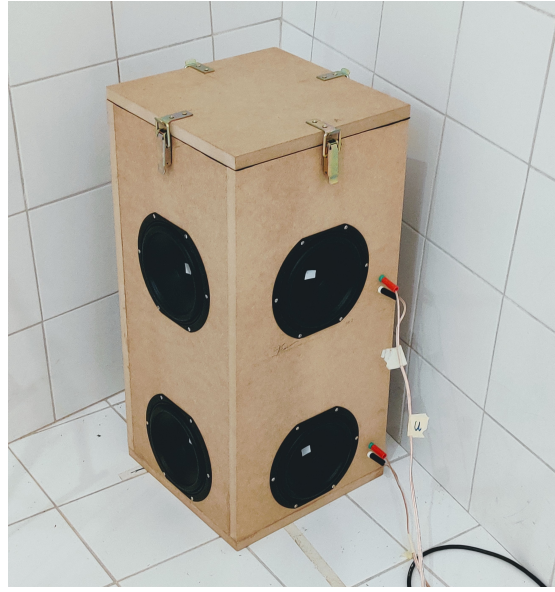


Figure 7.13 – The new prototype for the Electroacoustic Absorber

In this section, we will implement a 1-DOF model of the Electroacoustic Absorber to target a single mode of the room and try to see if we can reach its optimal resistance value (w.r.t the total area of absorption of 16 speakers from 2 absorbers). Technically, if the optimal resistance is reached, increasing or reducing this value will reduce the damping of the mode.

To test the capabilities and stability of the new prototype, we target the first two modes of the reverberation chamber which are both below 30 Hz (specifically, $f_1 \approx 20.6$ Hz and $f_2 \approx 26.7$ Hz. These two modes can be seen in Figure 7.14. From Figure 7.4, it should be estimated that with the same absorption area, these are two of the modes that require the lowest value of the optimal resistance.

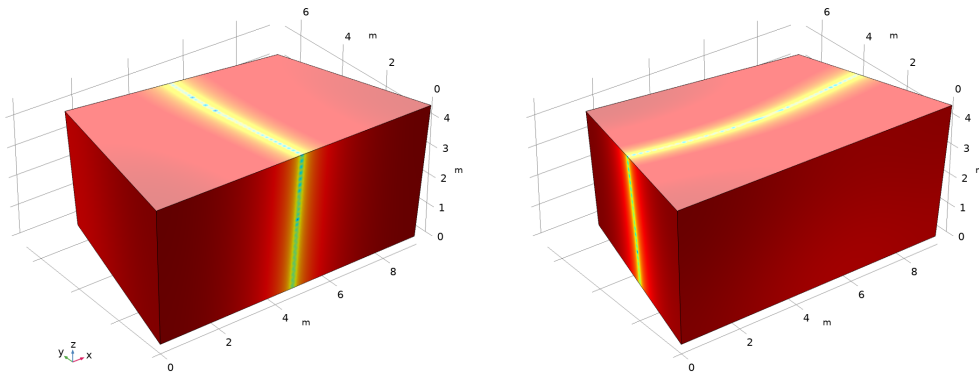


Figure 7.14 – The two targeted modes of the optimal resistance test.

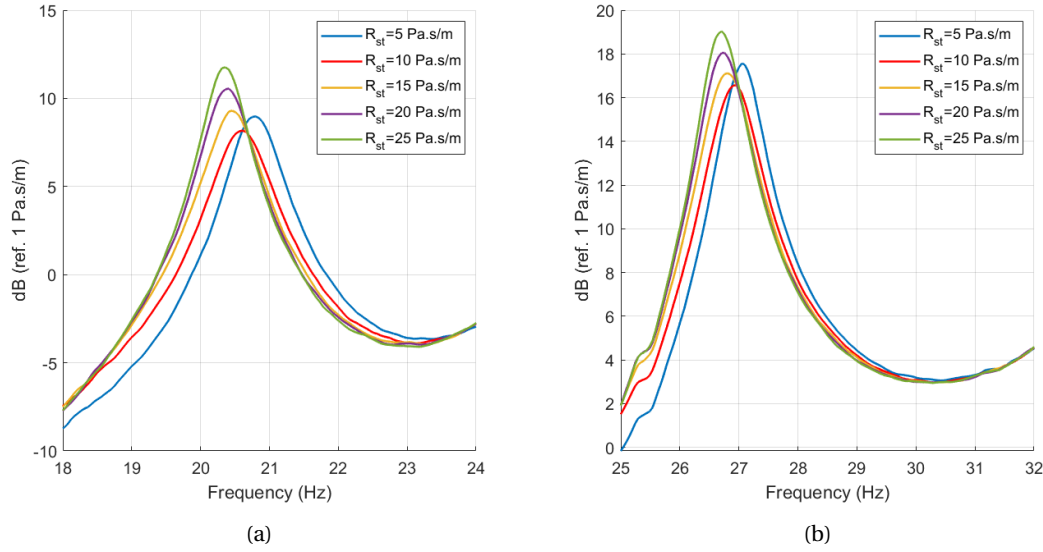


Figure 7.15 – Optimal specific acoustic impedance test for the two first modes of the reverberation chamber at a) 20.6 Hz and b) 26.7 Hz.

Using a narrow band impedance (so that the resistance could be set as low as possible) we position the center frequency of the target impedance at exactly the location of these two modes to observe the changes in its damping. Based on the values calculated from the FEM studies of the reverberation chamber in section 7.1.1, we set the range of R_{st} to be from 25 to 5 Pa.s/m. The bandwidth was set to be constant at only 5 Hz as it is the borderline value for $R_{st} = 5$ Pa.s/m in terms of stability. Figure 7.15 details the experimental results we get from sweeping this value at the location of each of the two modes.

It can be seen that by sweeping the value of R_{st} between 25 and 5 Pa.s/m, we have been able to observe the existence of the optimal resistance for the first two modes of the room. From reducing the resistance from 25 Pa.s/m, we first observe an increase in the damping of the targeted modes. This tendency continues down until the location of the optimal resistance. For both the 1st and 2nd mode, the optimal resistance can be seen to lie somewhere between 5 and 15 Pa.s/m. This is observed by noticing that when $R_{st} = 10$ Pa.s/m, the targeted mode has higher damping than both when $R_{st} = 5$ Pa.s/m and $R_{st} = 15$ Pa.s/m. This proves the existence of the optimal resistance which has been observed in simulation the reverberation chamber model in Section 7.1.1.

When compared to the specific value in Figure 7.4, at $S_{total} \approx 2500 \text{ cm}^2$, there are mismatches between the real situation and its numerical study. From the numerical studies, it seems like the optimal resistance value is at 15 Pa.s/m while for the 2nd mode, this value is at around 21 Pa.s/m. This once again highlights the fact that relying solely on the numerical studies to find optimal resistance could introduce significant errors to the estimation of the control. This is also one of the motivation behind the research in this chapter with the objective to fine tune

the control parameters based on real measurements.

7.2.3 Fine-tuning of optimal impedance using Sound field reconstruction

In Section 7.1.3, we have investigate some options for equalization metrics that are based on a large number of frequency responses within a domain in the room. From this analysis, two methods based on the relative moving standard deviation (from now, we refer to this method as **M1**) and the Relative Root mean square deviation (**M2**) are proposed to evaluate the equalization aspect of the active absorbers. It was then suggested from numerical studies that these two methods do have significant similarities in their evaluation results. Back then, the two methods were used to evaluate different combinations between the area of the absorbers and its corresponding target resistance. In real situation, most of the time, the area of absorbers are constrained. Furthermore, as explained in Section 7.1.1, a purely resistive impedance could not be implemented on the membrane with ease, especially when we try to achieve a very low value of R_{st} . This is why in this section, our study will concern the case where the total effective absorption area is fixed while the target 1-DOF impedance remains complex and can be adjusted using three deciding parameters: the resistance R_{st} , the bandwidth BW and the center frequency f_c .

For the 1-DOF resonator model of the target impedance, those three parameters are completely sufficient to adjust every aspect of the impedance. We recall the single resonator model and its control parameters, i.e, R_{st} , ν_1 and ν_2 from Eq. 6.11:

$$Z_{st}(\omega, \nu_1, R_{st}, \nu_2) = j\omega \frac{M_{ms}}{S_d \nu_1} + R_{st} + \frac{1}{j\omega S_d \nu_2 C_{mc}} \quad (7.4)$$

From this equation, the relation between the bandwidth and center frequency with the three parameters are derived below:

$$f_{center} = \frac{1}{2\pi \sqrt{M_{ms} C_{mc}}} \sqrt{\frac{\nu_1}{\nu_2}} \quad (7.5)$$

$$BW = \frac{1}{2\pi} \left(\frac{S_d \nu_1 R_{st}}{M_{ms}} \right) \quad (7.6)$$

From these two equations, we can derive the value of the direct control parameters ν_1 and ν_2

based on R_{st} , BW and f_c using:

$$\nu_1 = 2\pi \left(\frac{M_{ms}}{S_d R_{st}} \right) BW \quad (7.7)$$

$$\nu_2 = \left(\frac{1}{2\pi f_c} \right)^2 \frac{\nu_1}{M_{ms} C_{mc}} \quad (7.8)$$

With a fixed values of R_{st} , ν_1 and ν_2 , the target impedance can be fully determined using Eq. 7.4.

For this experiment, we want to test if sound field reconstruction could assist us in the design of active absorbers, i.e, giving us information regarding the design parameters of the control for the 1-DOF resonator case. The experiment set-up is as follow: Two prototypes of the Electroacoustic absorber are put in two corners of the reverberation chamber (Fig.7.16). As mentioned before, both absorbers are controlled by simulink real time through the use of Speedgoat I/O module. The controller transfer function Φ can be changed in real time by updating the values of R_{st} , ν_1 and ν_2 . A source to generate sound in the room is also controlled by simulink so that it can be turned on and off from the same platform. In front of the source is the velocimeter to measure the reference source velocity.

Inside the room are 15 microphones which positions are known. These microphones are spread randomly in the room. However, as the number of microphone is quite small in this case, in order to not leave out certain areas in the room, the microphone placement was randomized by Matlab to not leave out certain crucial area of the rectangular volume V_{rec} used for reconstruction. The microphone placement was also regulated by a single rule that at least 2/3 of the number of microphones should be located outside of V_{rec} and the rest inside the volume. Through initial tests, this has been shown to bring good balance to the reconstruction. The dimensions for V_{rec} is $5.5 \text{ m} \times 3.5 \text{ m} \times 1.5 \text{ m}$. The varying height inside the volume is chosen to be between 1 and 2.5m which fully covers both sitting and standing listening situation.

The experiment strategy is as follow: Various different settings of the 1-DOF impedance will be applied on the electroacoustic absorbers one by one. These settings are bounded either by some pre-acquired knowledge of the room or by limitation of the control regarding stability. For each setting of the impedance, the room responses of the 15 microphones in the room will be measured in order to be fed into the reconstruction framework. The framework then will reconstruct the sound field based on 12 out of the 15 microphones, the rest of them are kept as evaluation microphones so that the quality of the reconstruction can always be guaranteed. Using the signals from these 12 microphones, the responses of the $6 \times 6 \times 6$ points that scan



Figure 7.16 – Experiment set-up with 2 Electroacoustic Absorbers at the corners and 15 randomly placed microphones in the room.

V_{rec} can be recovered and used to calculate the metrics introduced in Section 7.1.3.

One of the advantages of the experiment set up is that through the use of Simulink real-time and Matlab, most of the processes can be programmed to be automated with very little supervision needed. As both the target impedance and the source can be controlled in real time, all the different settings can be arranged in series with a short silent interval of the source in between to allow the reverberation from the previous setting to die out before starting the new one. The measurements of the microphones are kept to be continuously recording throughout the whole experiment. After the entire measurements are finished, the measurement data is sent directly to Matlab and cut into individual segments. Each individual segment represents a single target impedance setting.

In this study, we target the frequency range between 18 Hz and 48 Hz where around 7 modes of the room reside. The objective is to sweep through different target impedance curves and look for the one that will bring us the best equalization performance based on the metrics we introduced previously. In this case, we tested with a total of 48 different settings for the target impedance, with different combination of the resistance $R_{st} = 5, 10, 15$ or 20 Pa.s/m , bandwidth $BW = 5, 10, 15$, or 20 Hz and center frequency $f_c = 25, 30$ or 35 Hz . The ranges for these values are based on some brief initial testing. As mentioned before, some of the combination could not be realized due to technical instability of the control. These combinations are listed in Table 7.2.

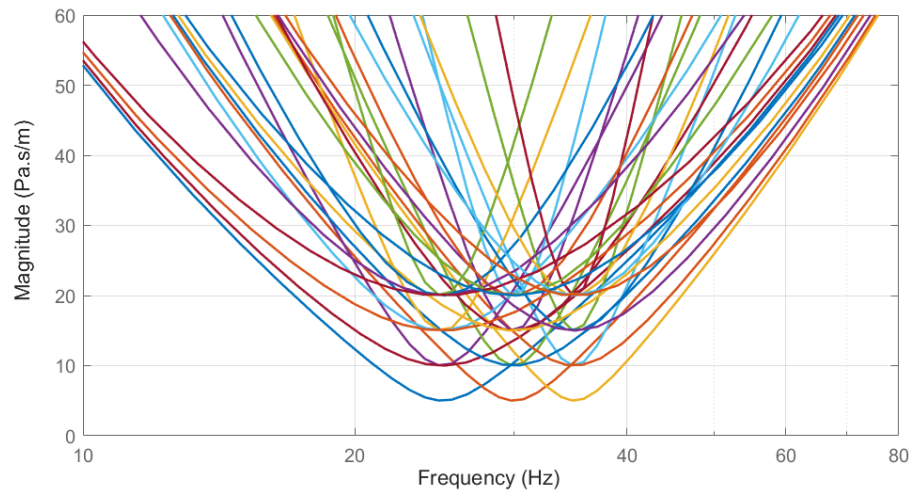


Figure 7.17 – The 30 stable scenarios for the Target Impedance of the electroacoustic absorbers.

Out of the 48 possible combinations of the controls, 18 of them are unstable and hence could not be analyzed. This leaves us with 30 different combinations for further study. The 30 different specific acoustic impedance curves are presented altogether in Figure 7.17.

Using sound field reconstruction for each of these cases and calculate the two metrics **M1** and **M2** for the frequency band of [18 Hz , 48 Hz] could provide us with crucial information regarding the efficiency of each of these settings. As a first evaluation, the results of all 30 settings are featured in Figure 7.18.

The first observation that can be made is that once again, we can see that in general, **M1** and **M2** returns quite similar evaluation. Certain differences arises more for configurations

Table 7.2 – Combination of design paramters that results in instability for the control

| # | R_{st} (Pa.s/m) | BW (Hz) | f_c (Hz) |
|---|----------------------|--------------|---------------|
| 1 | 5 | 10 | 25 |
| 2 | 5 | 10 | 30 |
| 3 | 5 | 10 | 35 |
| 4 | 5 | 15 | 25 |
| 5 | 5 | 15 | 30 |
| 6 | 5 | 15 | 35 |
| 7 | 5 | 20 | 25 |
| 8 | 5 | 20 | 30 |
| 9 | 5 | 20 | 35 |

| # | R_{st} (Pa.s/m) | BW (Hz) | f_c (Hz) |
|----|----------------------|--------------|---------------|
| 10 | 10 | 15 | 25 |
| 11 | 10 | 15 | 30 |
| 12 | 10 | 15 | 35 |
| 13 | 10 | 20 | 25 |
| 14 | 10 | 20 | 30 |
| 15 | 10 | 20 | 35 |
| 16 | 15 | 20 | 25 |
| 17 | 15 | 20 | 30 |
| 18 | 15 | 20 | 35 |

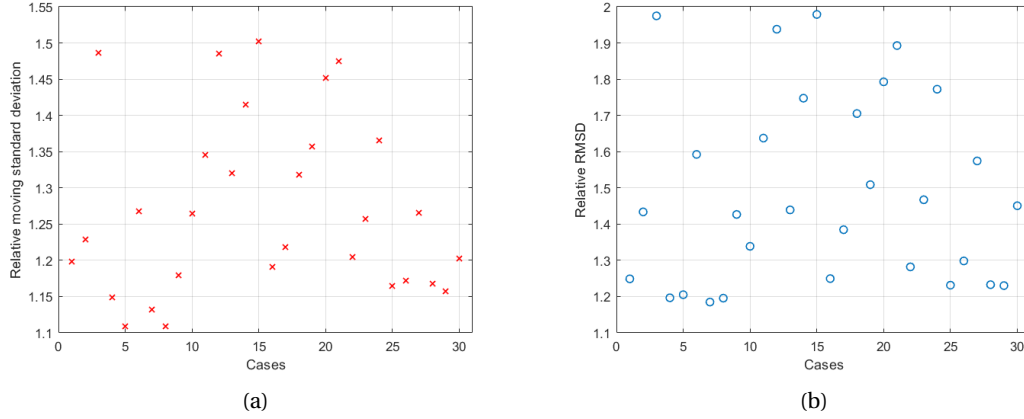


Figure 7.18 – The equalizing performance of all 30 stable cases of the target impedance in studies evaluated by a) metric **M1** and b) metric **M2**.

with a higher deviation value but for the good configurations, the two methods more than usual, agree in terms of results. This initial evaluation will first help us eliminate most of the non-optimal settings. Recall that in the end, this is a 4-D function of evaluation as there are three variable parameters that can affect the evaluation result. Hence, in order to understand more clearly the relationship that each parameter can have with respect to the results, we suggests multiple ways to express this result graphically.

The first analysis is with regards to the center frequency. There are 3 different choices for the location of the peak of the target impedance, at 25 Hz, 30 Hz and 35 Hz. In Figure 7.19, we plot the effect that the center frequency has on the equalization performance of the control based on method **M2**. The solid lines refer to configurations with the a fixed value for R_{st} and BW .

It can be clearly seen from the analysis that when the value of R_{st} and BW are fixed, moving the center frequency of the impedance does affect the overall result. For all the cases within our scope of analysis, the performance deteriorates when the center frequency is increased to 35 Hz. Between 25 Hz and 30 Hz, we can see two different trends. For some of the most optimal configurations (i.e, low value of RMSD), the performance stays quite stable for both 25 Hz and 30 Hz, while for the rest, placing the peak at 25 Hz results in better performance than at 30 Hz.

Following the same procedure, we analyze the effects bandwidth has on the performance of the absorbers. The results can be seen in Figure 7.20. As previously explained, some of the configurations of the control could not achieve stability. This is mostly due to the combination between the bandwidth and the specific acoustic resistance R_{st} of the control. The lower we force R_{st} to go, the narrower the bandwidth has to be to avoid instability. In this case, we have an intriguing coincidence between them such that when $R_{st} = 5 \text{ Pa.s/m}$, the bandwidth value (from our list of chosen bandwidth) could not be increased to higher than 5 Hz. The same

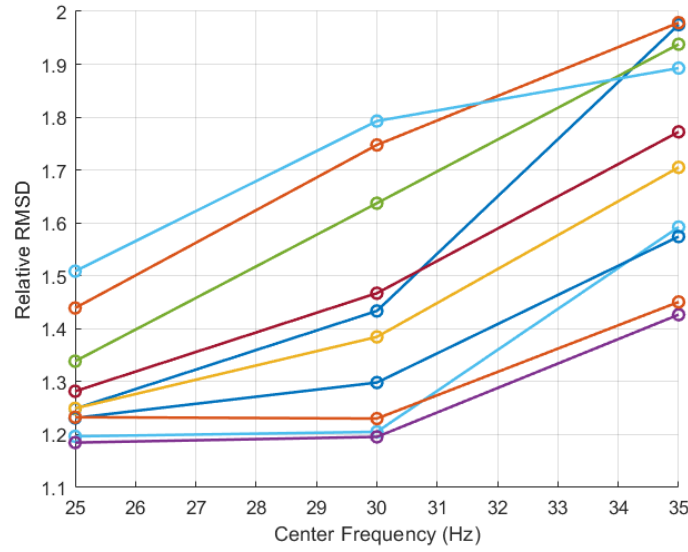


Figure 7.19 – The effects of the center frequency on different settings of the control. Connected points represent settings with the same R_{st} and BW with varying f_c .

goes for $R_{st} = 10, 15$ and 20 Pa.s/m, the bandwidth respectively, could not increase to higher than $BW = 10, 15$ and 20 Hz. As mentioned, this is purely a coincidence in terms of value and should not correspond to any discernible relationship. This limitation due to instability can be observe in Fig. 7.20 as not all the configuration can go through with all available bandwidth from 5 Hz to 20 Hz.

From the analysis, it can be seen that for all the configuration under our scope of studies, the bandwidth generally should be increased as much as possible as the performance tends to increase with a wider bandwidth. However, this statement should not be generalized to other studies as depending on the relations between the optimal resistance for the modes, there might exist an optimal bandwidth that drives the real part of the impedance in such a way that it goes past multiple optimal resistance of the modes and hence results in a better performance. For this particular case at very low frequencies, we acknowledge the fact that when bounded by instability, the performance in this case increases along with bandwidth.

Finally, we discuss the effects of the most important factor of the control parameters, the specific acoustic resistance R_{st} . For R_{st} , a plot similar to ones in Fig. 7.19 and 7.20 would not be that useful as it is entangled by the mostly uncorrelated relationship between the center frequency and the bandwidth. A better way to present the results would be to study the relationship between R_{st} and the bandwidth at each center frequency. Figure 7.21 plots three different color maps produced by Matlab for each of the center frequencies of the impedance. Due to the coincidence between the bandwidth and R_{st} , the top left half of graph could not be realized due to instability.

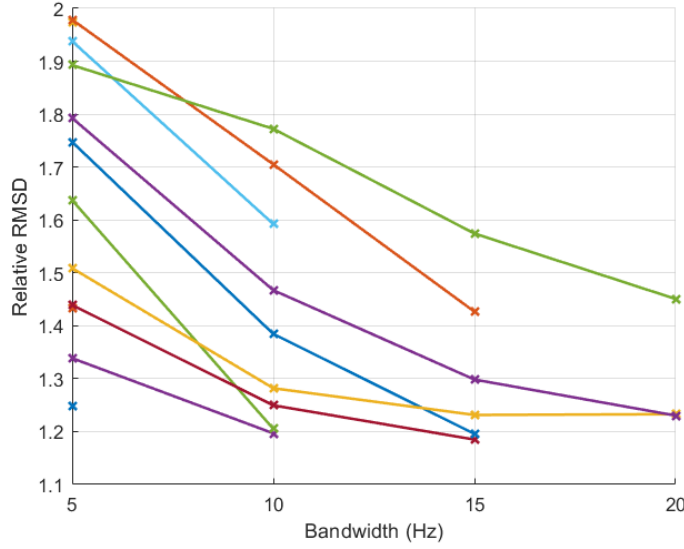


Figure 7.20 – The effects of the bandwidth on different settings of the control. Connected points represent settings with the same R_{st} and f_c with varying BW.

From the color maps, it could be seen that at each center frequency, there is a certain optimal regions that can bring out the best performance of the absorber. For example, for $f_c = 25$ Hz and 30 Hz, we can see that R_{st} between 10 and 15 Pa.s/m seems to produce the best performance based on **M2** criteria. Furthermore, as observed from the previous center frequency analysis, the overall performance at 25 Hz is slightly better than that at 30 Hz. For the case when $f_c = 35$ Hz, the optimal region, though exists, performs much less effective when compared to the previous two center frequencies. This results agree with the overall behavior of the optimal resistance for the first few modes of the room, where the first two modes (at 20.4 and 26.7 Hz) have much lower optimal value of R_{st} than the modes that follow. Placing the center frequency quite far from the right of them would mean that the real part of the impedance at these two frequencies will be too high to be effective in damping these modes.

From these analysis, we can see that taking R_{st} between 10 and 15 Pa.s/m and placing them somewhere in between 25 Hz and 30 Hz will produce an effective performance of the equalization. The exact same analysis procedure is performed using the **M1** methods which points to the same conclusion (see Appendix B).

To have a reference of the efficacy of the performance, for the situation of the room without any control, the average value for the relative RMSD is 1.99 which is much higher than the optimal cases according to the metrics. Figure 7.22 plotted the response at a particular point in the room for two different settings of the target impedance. Case A is one of the non-optimal settings for the control with a quite low value of RMSD = 1.34 while Case B (with $R_{st} = 15$ Pa/m, BW = 15 Hz and $f_c = 25$ Hz) is the optimal out of the 30 settings according to our metrics with

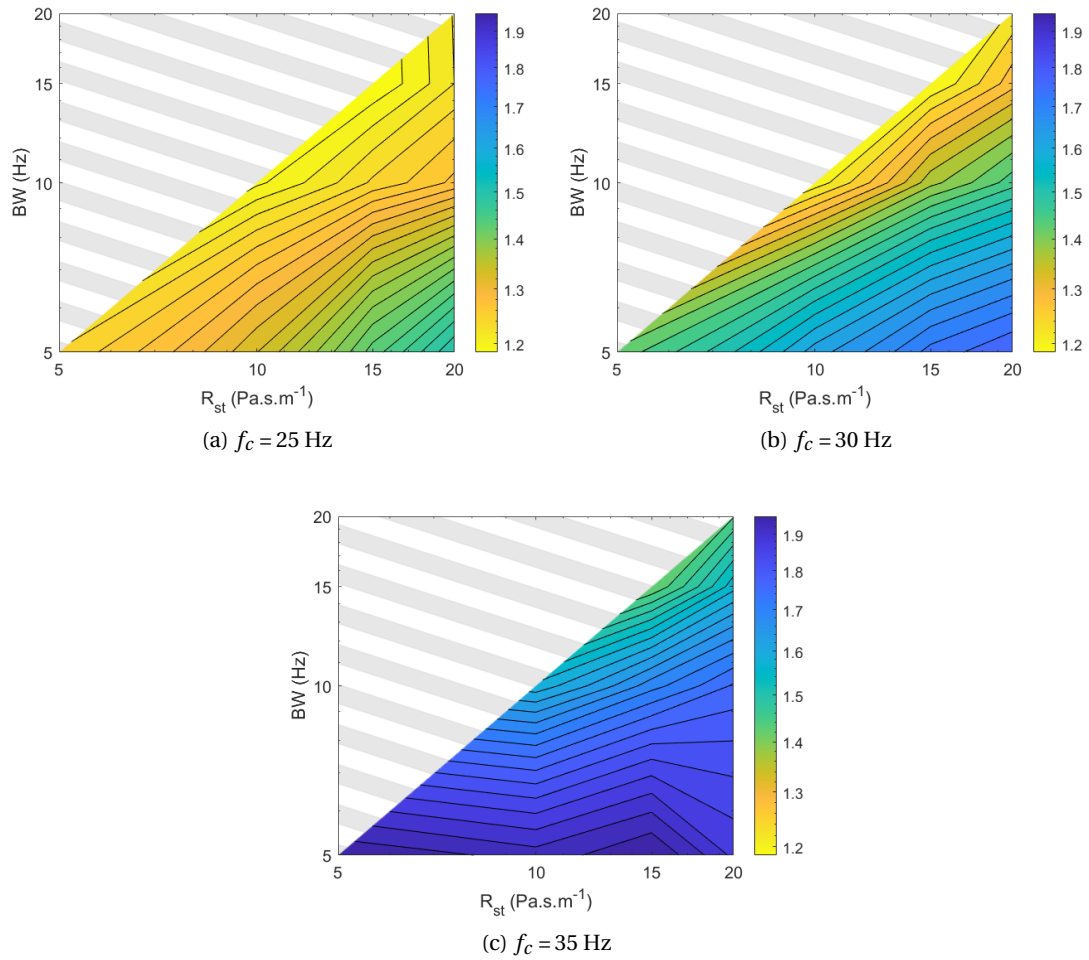


Figure 7.21 – Evaluation of metric **M2** with respect to R_{st} and bandwidth at different center frequency of target impedance: a) 25 Hz b) 30 Hz, c) 35 Hz. The top left region represents area that couldn't be realized at the membrane due to instability.

RMSD = 1.19.

We can see that there is a clear universal reduction in fluctuations of the pressure across the entire frequency band of analysis for both case A and B compared to the case without any control. Case B with the more optimal settings also performs better in equalizing room modes than case A at 3 different modes. We can also observe that the target impedance for case B is able to lower the sound pressure for the 1st mode by an impressive value of around 20 dB. Finally, as there are only two Electroacoustic Absorber in this case but 4 corners to place them in the room, we further tested with a different combination of corners in the room and arrived at the same conclusion for the range of R_{st} . There is predicted to be more difference, however, if the analysis goes into the higher order modes where the nodal lines could be closer to some particular corners of the room.

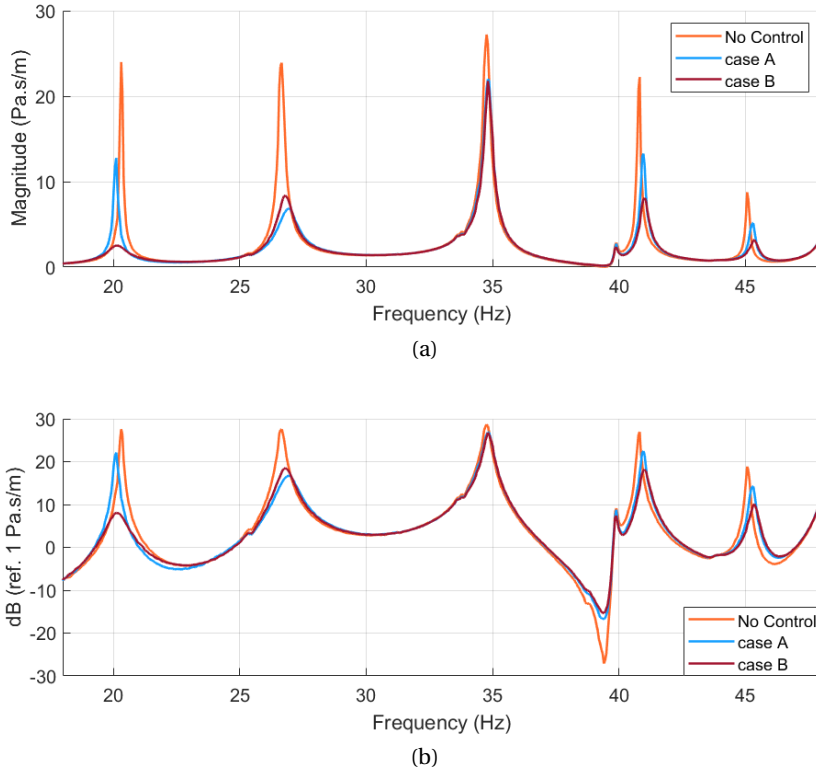


Figure 7.22 – Individual response comparison between two different settings of the controller for target impedance in real setup as compared to the room response without any active control presented in a) magnitude and b) decibel.

The analysis above shows that by using the metrics we proposed in Section 7.1.3 in combination with an automated procedure of sweeping through different combination of the control parameters, we can not only locate the optimal settings within the generated set of study but further more acquire precious information regarding the relationship between the control parameters and the equalization performance of the absorbers.

This study can be apply to any targeted area of the room depending on custom preference. Thanks to the robustness of the reconstruction framework that we have, the un-supervised automated reconstruction process all returned highly accurate results. Using the 12 microphones for reconstruction and checking the validity using the other 3, we achieve a universally high reconstruction results of the sound field between 18 and 48 Hz. The average correlation value of $COR\%$ is reliably high at 99.1 % with a stable standard deviation of 0.5%. As the upper frequency in this case is less than 50 Hz, the number of microphones have been significantly reduced as compared to the cases in Chapter 5. Even in this very study of this section, we could possibly reduce the number of microphones to just 9 microphones with an acceptable penalty for the correlation value where average $COR\%$ becomes 98.7 % with a standard deviation of 0.8 %.

The proposed method from this section could open doors to multiple future application. One of the possibility is regarding the design of the multi-degree of freedom resonator model. There are multiple strategies that could be further studied. The first potential direction is to expand the same study that has been done here to the second resonators by fixing the first resonator using the optimal parameters above and then try sweeping through different settings of the second one. Another idea is to apply the knowledge from the domain of design of experiments and actively sweeping through different parameters for multiple resonators at the same time and locate the optimal that usually is not contained in the original set. This would require a more practical set of combination between these parameters for the solution to converge. It is also unsure how likely this type of design could reach for an unstable solution. Both method would of course, require an increase in the number of microphones in the room as the frequency range moves up. Thanks to both the automation processing and the robustness of the sound field reconstruction framework, the measurements and post processing process could remain nearly the same without any major modification.

7.3 Conclusion

In this chapter, we have analyzed a more complex use of the sound field reconstruction framework in the field of modal equalization of sound field in a room at low frequencies. We have first examined the design of the Electroacoustic Absorber and its foundation which is based on the concept of target impedance. We also briefly reviewed past optimization attempts based on a vast simulation data from a generic set of 3 different room sizes and geometries. This process, in itself, proved to be difficult and inconvenience to be replicated using measurements from real rooms. Driven by this motivation, we proceed to look for different ways to assess the equalization of the sound field which takes into account the performance for a large number of locations within a frequency range instead of individual modal metrics. From this, two different equalization qualification metrics were suggested, both based on the assessment of how much the Frequency responses deviates across the frequency range. The two metrics both show promising and similar potential in assessing the equalizing effects of active absorbers.

We then have taken a step further by using these metrics as assessment tools to modify and fine tune the design of the impedance of a single resonator model for the Electroacoustic Absorber. The process have been made to be as automated as possible where a large number of different settings for the loudspeaker membrane of the EAs are tested continuously and in order. Thanks to the sound field reconstruction framework which has been tested in Chapter 5 and 6, the sound field of the rectangular sub-space of the room is also continuously reconstructed for each and every impedance settings. Using the aforementioned metrics, the control parameters of the single resonator can be tuned to achieve the highest performance. Moreover, through this parametric sweep studies, a deeper understanding of design parameters such as bandwidth and center frequency can also be acquired.

The tuning process in this chapter open doors to multiple promising application in the future, specifically the design of a multi-degree-of-freedom target impedance model based on real measurements or further more, a different impedance shaping methods that could make use of the sound field reconstruction evaluations.

8 Conclusion and perspectives

The work included in this thesis is intended towards sound field reconstruction in a room at low frequencies and its potential application in the field of room modal equalization. Most of the analysis in the thesis are focused on the practical aspects of the framework and its potential application. At the end, we have been able to design a reconstruction framework that performs with high recovery precision of the spatial responses in the room using a small set of actual measurements. This allows us to extend our understanding of the room acoustic behavior at low frequencies through a different scope. Through the research in this thesis, we have also been able to show a small part of the vast potentials that the sound field reconstruction framework can have on multidisciplinary, with a special focus on the field of room modes equalization where some novel metrics and procedures have been proposed to assist with the design of active absorbers. Below are the summary of the research that has been done as well as our perspective and discussion on future directions in multiple domain.

Sound field reconstruction framework

In this thesis, we have provided the sparse analysis on the governing equation for the acoustics of the room at low frequencies. Through these analysis and derivations, we have been able to come up with a reconstruction framework that could perform with high accuracy using a limited set of measurements in a non-rectangular room. We have also analyzed the performance of different modal analysis methods, in this case SOMP and RFP. Due to the inherent robustness of the matching pursuit regime, SOMP tends to perform better than RFP when there is little to no supervision. RFP on the other hand, could converge much quicker than SOMP however it requires some a priori information of the room to converge to the correct results. Looking further, there are still many other different methods that could be potentially investigated for this process such as vector fitting (which could perform as quick as RFP), Multiple Signal Classification (MUSIC) and other matching pursuit method such as CoSAMP. These could all be investigated in the future based on two main criteria which is accuracy and cost-effective.

Another aspect of the framework that could be further investigate is the mode shapes approximation process and the incorporated spherical sampling. This process is directly linked with the accuracy of the reconstruction results, especially in cases where the number of microphones are extremely low. We have suggested the use of an overly sampled set of plane waves that are larger than the number of measurement points in these situation. However, more detailed research regarding regularization limit should be looked into in the future. Lastly, a more adaptive plane waves sampling process could be investigated to make the most out of each projection of the mode shape function on individual plane waves.

Sound field reconstruction performance

In this thesis, we have proposed multiple methods to analyze the reconstruction algorithm. These ranges from the analysis regarding the graphical representation of modal behaviors of the room to average spatial correlation analysis. Moreover we also investigated the relationship between the reconstruction accuracy and the number of input measurement points. There are still other aspects of the reconstruction that could be investigated further, such as its performance closer to the walls of the room which could open doors to potential application in wall damping analysis. Furthermore, an interesting aspect that could be investigated is how the reconstruction performs in the case of an unbalanced distributed placement of microphones in the room. For example, how well would the reconstruction perform for the sound field on the left side of the room when most microphones are on the right. This could also be extended to the comparison between interpolation with extrapolation performance of the framework.

We have also been able to show the high precision performance of the reconstruction for the sound field of the room under the influence of enclosed active elements such as the electroacoustic absorbers. The results show that the framework is robust and can be used in a diverse field of boundary conditions.

Sound field reconstruction application in Room modes equalization

In this research, we have been able to show a unique application of the sound field reconstruction framework. When a spatial reconstruction of a space could be recovered, the most direct application of the framework is in the field of sound reproduction as well as auralization. However, we would like to show that the reconstructed sound field can also be used effectively in the field of modal equalization. In order to do this, some novel metrics have been developed to take into account the equalization effects on the frequency responses that could be reconstructed using our framework. These metrics have shown promising results in analyzing the flatness of a frequency response. However, as discussed in Section 7.1.3, there are still many other ways to define different metrics based on the application, such as metrics based on equal distribution of energy for different bandwidth of analysis, or metrics based

Conclusion and perspectives

on the dynamic of the frequency responses... Our suggested metrics are just two of the many different possible ways that can be investigated in future research.

Throughout Section 7.2, we have been able to suggest a method that can make use of the reconstruction results of the sound field in the room to benefit the design of the active electroacoustic absorbers. Within the sweeping studies of the control parameters, we have been able to locate the most optimal settings that could return a high equalization performance. This proves that the reconstruction framework can truly be beneficial to the domain of room modes equalization. The research in this thesis have only focused on experimenting the idea on the single resonator model of the target impedance. As explained previously, the analysis in Section 7.2 could be extended to a multi-degree-of-freedom resonator model of the target impedance, where a higher number of design parameters could be taken into consideration. This would certainly increase the complexities of the study but also poses very interesting prospects. Future research could focus on this idea in order to develop a broadband active absorbers that could adapt to the actual room situation based on the data from the reconstructed sound field of the room.

A Additional concepts and formula

Rational fraction polynomial and global curve fitting

Considering the case of a linear system that is of the second dynamical order, its frequency responses can be expressed as a ratio of two polynomials in the Laplace domain ($s = j\omega$) using the Rational fraction polynomial form:

$$H(\omega) = \frac{\sum_{i=0}^m a_i s^i}{\sum_{k=0}^n b_k s^k}, a_i \text{ and } b_k \in \mathbb{R}. \quad (\text{A.1})$$

Moreover, in cases where the system is resonant, i.e, the response is governed by its resonances, the Frequency response can be reformulated using the partial fractional form to put emphasis on the poles of the system.

$$H(\omega) = \sum_{k=0}^{n/2} \left[\frac{r_k}{s - p_k} + \frac{r_k^*}{s - p_k^*} \right] \quad (\text{A.2})$$

where $p_k = j\omega_k - \delta_k$ is the k^{th} pole of the system and r_k is the corresponding residue.

The curve fitting procedure focuses on minimizing the squared error J between the analytical and measured response computed at each and every frequency bin e_i :

$$J = \sum_{i=1}^L e_i^* e_i \quad (\text{A.3})$$

where L is the length of the frequency vector. Assuming now that multiple different frequency response functions of the system were measured, they all will contain these same inherent poles and hence the denominator of each and every measurements shall contain the same characteristic polynomial. This claim can be proved to be true for most cases since for a resonant-based system such as the one in room acoustics, the modal frequencies and damping are always the same regardless of where they are measured within the room. This further simplifies the curve fitting procedure which allows for a global curve fitting [52] of the entire set of measurements to recover the parameters in Eq. (A.1) or Eq. (A.2).

To avoid problem arising from ill-conditioning, the frequency response function can be reformulated using orthogonal polynomials only in the positive domain of the frequency axis through the use of the Forsythe method[43]. This greatly simplifies the computation of the problem and although the final result is expressed in the orthogonal function expansion form, it is not difficult to trace back to the form in Eq. (A.2) for original information of the poles and residues. This method, detailed in Ref. [98], is non-iterative and efficiently fast. Furthermore, by reasonably choosing m and n , it can compensate the effects created by out-of-band modes and hence, reduce the fitting error by a significant amount. As any other curve fitting method in the frequency domain, it suffers from over-fitting as well as from the lack of frequency resolution. This could be further improved using the Interpolated DFT detailed in [33, 104, 14].

B Additional figures

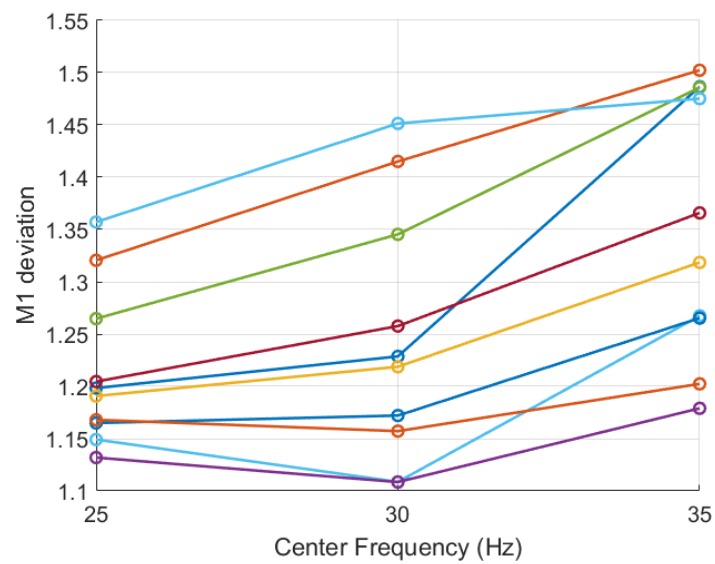


Figure B.1 – The effects of the center frequency on different settings of the control according to metric **M1**

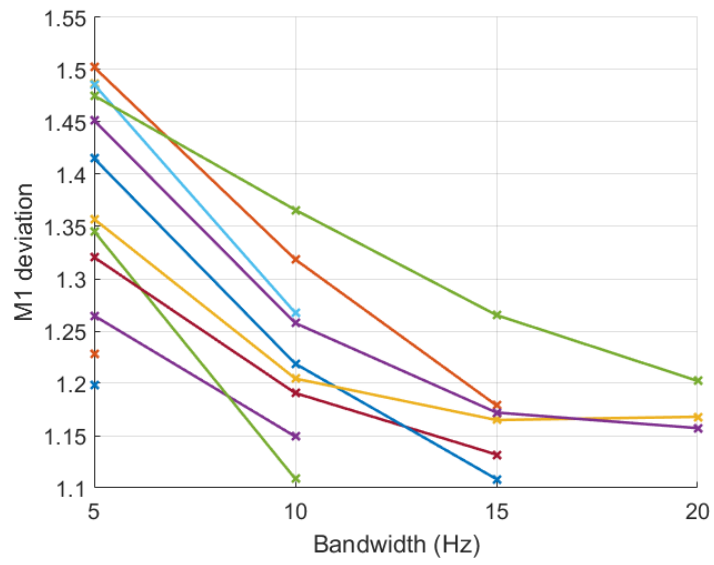


Figure B.2 – The effects of the bandwidth on different settings of the control according to metric **M1**

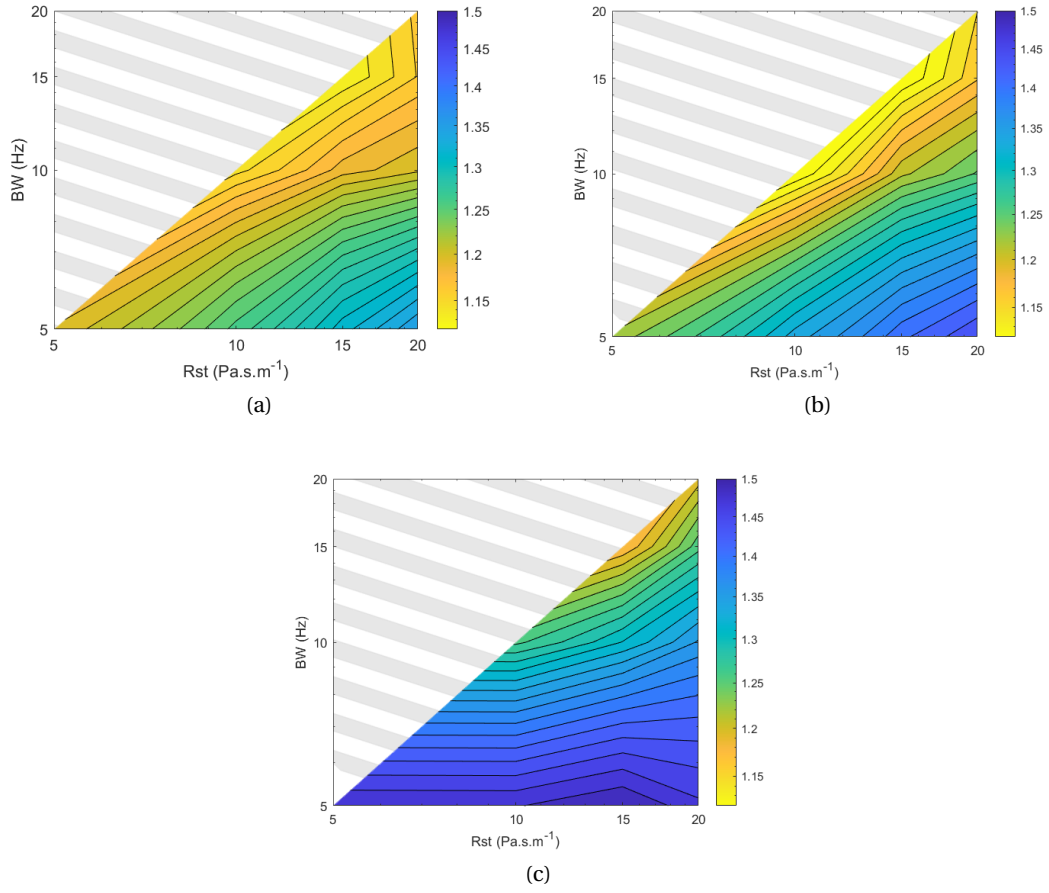


Figure B.3 – Evaluation of metric **M1** with respect to R_{st} and bandwidth at different center frequency of target impedance. The top left region represents area that couldn't be realized at the membrane due to instability.

Bibliography

- [1] T. Ajdler, L. Sbaiz, and M. Vetterli. “The Plenacoustic Function and Its Sampling”. In: *IEEE Transactions on Signal Processing* 54.10 (Oct. 2006), pp. 3790–3804. ISSN: 1053-587X. DOI: 10.1109/TSP.2006.879280.
- [2] Thibaut Ajdler. “The plenacoustic function and its applications”. PhD thesis. 2006.
- [3] Ali N. Akansu and Richard A. Haddad. *Multiresolution signal decomposition: transforms, subbands, and wavelets*. Academic Press, 2001.
- [4] N. Antonello et al. “Joint Acoustic Localization and Dereverberation Through Plane Wave Decomposition and Sparse Regularization”. In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 27.12 (Dec. 2019), pp. 1893–1905. DOI: 10.1109/TASLP.2019.2933047.
- [5] Jorge Arenas and Malcolm Crocker. “Recent Trends in Porous Sound-Absorbing Materials”. In: *Sound & vibration* 44 (July 2010), pp. 12–17.
- [6] Balazs Bank. “Warped IIR filter design with custom warping profiles and its application to room response modeling and equalization”. In: *130th Audio Engineering Society Convention 2011* 2 (Jan. 2011), pp. 13–16.
- [7] Alex H. Barnett and Timo Betcke. “Stability and convergence of the method of fundamental solutions for Helmholtz problems on analytic domains”. In: *J. Comput. Physics* 227 (2008), pp. 7003–7026.
- [8] Julius S. Bendat and Allan G. Piersol. *Engineering applications of correlation and spectral analysis*. J. Wiley, 1993.
- [9] Leo L. Beranek. *Acoustic measurements*. Wiley, 1965.
- [10] F. Bergeaud and S. Mallat. “Matching pursuit of images”. In: *Proceedings., International Conference on Image Processing*. Vol. 1. 1995, 53–56 vol.1. DOI: 10.1109/ICIP.1995.529037.
- [11] Jeffrey Blanchard, Coralia Cartis, and Jared Tanner. “Compressed Sensing: How Sharp Is the Restricted Isometry Property?” In: *SIAM Review* 53 (Apr. 2010). DOI: 10.1137/090748160.
- [12] Liefeng Bo, Xiaofeng Ren, and Dieter Fox. “Hierarchical Matching Pursuit for Image Classification: Architecture and Fast Algorithms”. In: (May 2012).

-
- [13] Richard H. Bolt. “Normal Modes of Vibration in Room Acoustics: Experimental Investigations in Nonrectangular Enclosures”. In: *The Journal of the Acoustical Society of America* 11.2 (1939), pp. 184–197. DOI: 10.1121/1.1916022. eprint: <https://doi.org/10.1121/1.1916022>. URL: <https://doi.org/10.1121/1.1916022>.
 - [14] Józef Borkowski, Dariusz Kania, and Janusz Mroczka. “Interpolated-DFT-Based Fast and Accurate Frequency Estimation for the Control of Power”. In: *IEEE Transactions on Industrial Electronics* 61.12 (2014), pp. 7026–7034. DOI: 10.1109/TIE.2014.2316225.
 - [15] Romain Boulandet, Etienne Rivet, and Hervé Lissek. “Sensorless Electroacoustic Absorbers Through Synthesized Impedance Control for Damping Low-Frequency Modes in Cavities”. In: *Acta Acustica united with Acustica* 102 (July 2016). DOI: 10.3813/AAA.918986.
 - [16] Romain Boulandet, Etienne Rivet, and Hervé Lissek. “Sensorless Electroacoustic Absorbers Through Synthesized Impedance Control for Damping Low-Frequency Modes in Cavities”. In: *Acta Acustica united with Acustica* 102.4 (2016), pp. 696–704. ISSN: 1610-1928. DOI: doi:10.3813/AAA.918986. URL: <https://www.ingentaconnect.com/content/dav/aaua/2016/00000102/00000004/art00010>.
 - [17] Lars-Johan Brannmark, Adrian Bahne, and Anders Ahlen. “Improved loudspeaker-room equalization using multiple loudspeakers and MIMO feedforward control”. In: Mar. 2012, pp. 237–240. ISBN: 978-1-4673-0045-2. DOI: 10.1109/ICASSP.2012.6287861.
 - [18] Buffa, Annalisa and Monk, Peter. “Error estimates for the Ultra Weak Variational Formulation of the Helmholtz equation”. In: *ESAIM: M2AN* 42.6 (2008), pp. 925–940. DOI: 10.1051/m2an:2008033. URL: <https://doi.org/10.1051/m2an:2008033>.
 - [19] T. Tony Cai and Lie Wang. “Orthogonal Matching Pursuit for Sparse Signal Recovery With Noise”. In: *IEEE Transactions on Information Theory* 57.7 (2011), pp. 4680–4688. DOI: 10.1109/TIT.2011.2146090.
 - [20] Emmanuel J. Candes and Michael B. Wakin. “An Introduction To Compressive Sampling”. In: *IEEE Signal Processing Magazine* 25.2 (2008), pp. 21–30. DOI: 10.1109/MSP.2007.914731.
 - [21] Emmanuel Candès, Justin Romberg, and Terence Tao. “Stable Signal Recovery from Incomplete and Inaccurate Measurements”. In: *Communications on Pure and Applied Mathematics* 59 (Aug. 2006). DOI: 10.1002/cpa.20124.
 - [22] Emmanuel J. Candès. “The restricted isometry property and its implications for compressed sensing”. In: *Comptes Rendus Mathématique* 346.9 (2008), pp. 589–592. ISSN: 1631-073X. DOI: <https://doi.org/10.1016/j.crma.2008.03.014>. URL: <https://www.sciencedirect.com/science/article/pii/S1631073X08000964>.
 - [23] Stefania Cecchi, Alberto Carini, and Sascha Spors. “Room Response Equalization—A Review”. In: *Applied Sciences* 8 (Dec. 2017), p. 16. DOI: 10.3390/app8010016.

BIBLIOGRAPHY

- [24] Gilles Chardon, Albert Cohen, and Laurent Daudet. “Sampling and reconstruction of solutions to the Helmholtz equation”. In: *Sampling Theory in Signal and Image Processing* 13 (2014), pp. 67–90. URL: <https://hal.archives-ouvertes.fr/hal-01350609>.
- [25] Gilles Chardon and Laurent Daudet. “Optimal subsampling of multichannel damped sinusoids”. In: *2010 IEEE Sensor Array and Multichannel Signal Processing Workshop* (2010). DOI: 10.1109/sam.2010.5606750.
- [26] Nazli Che Din et al. “A Practical Application of Ensemble Averaged Surface Normal Impedance Measured In-Situ”. In: *International Journal of Automotive and Mechanical Engineering* 7 (June 2013), pp. 1074–1085. DOI: 10.15282/ijame.7.2012.22.0087.
- [27] Shaobing Chen et al. “Basis Pursuit”. In: (Mar. 1996).
- [28] Lothar Cremer and Müller Helmut A. *Principles and applications of room acoustics*. Applied Science, 1982.
- [29] Elaine Crespo Marques et al. “A Review of Sparse Recovery Algorithms”. In: *IEEE Access* 7 (2019), pp. 1300–1322. DOI: 10.1109/ACCESS.2018.2886471.
- [30] George B. Dantzig. *Reminiscences about the Origins of Linear Programming*. Defense Technical Information Center, 1981.
- [31] Mark Davenport et al. “Introduction to compressed sensing”. In: *Preprint* 93 (Jan. 2012). DOI: 10.1017/CBO9780511794308.002.
- [32] Geoffrey M. Davis, Stephane G. Mallat, and Zhifeng Zhang. “Adaptive time-frequency decompositions”. In: *Optical Engineering* 33.7 (1994), pp. 2183–2191. DOI: 10.1117/12.173207. URL: <https://doi.org/10.1117/12.173207>.
- [33] Asja Derviškadić, Paolo Romano, and Mario Paolone. “Iterative-Interpolated DFT for Synchrophasor Estimation: A Single Algorithm for P- and M-Class Compliant PMUs”. In: *IEEE Transactions on Instrumentation and Measurement* 67.3 (2018), pp. 547–558. DOI: 10.1109/TIM.2017.2779378.
- [34] D.L. Donoho. “Compressed sensing”. In: *IEEE Transactions on Information Theory* 52.4 (2006), pp. 1289–1306. DOI: 10.1109/TIT.2006.871582.
- [35] David L. Donoho et al. “Sparse Solution of Underdetermined Systems of Linear Equations by Stagewise Orthogonal Matching Pursuit”. In: *IEEE Transactions on Information Theory* 58.2 (2012), pp. 1094–1121. DOI: 10.1109/TIT.2011.2173241.
- [36] Jeffrey Dosch, Daniel Inman, and Ephraim Garcia. “A Self-Sensing Piezoelectric Actuator for Collocated Control”. In: *Journal of Intelligent Material Systems and Structures* 3 (Jan. 1992), pp. 166–185. DOI: 10.1177/1045389X9200300109.
- [37] Jean-Baptiste Dupont and Marie-Annick Galland. In: *Active absorption to reduce the noise transmitted out of an enclosure*. 2009, pp. 142–152.
- [38] Samuel Dupont, Manuel Melon, and Alain Berry. “Absorption measurement of acoustic materials using a spherical microphone array”. In: *The Journal of the Acoustical Society of America* 145.3 (2019), pp. 1896–1896. DOI: 10.1121/1.5101875. eprint: <https://doi.org/10.1121/1.5101875>. URL: <https://doi.org/10.1121/1.5101875>.

-
- [39] Armin Eftekhari et al. "The restricted isometry property for random block diagonal matrices". In: *Applied and Computational Harmonic Analysis* 38.1 (2015), pp. 1–31. ISSN: 1063-5203. DOI: <https://doi.org/10.1016/j.acha.2014.02.001>. URL: <https://www.sciencedirect.com/science/article/pii/S1063520314000220>.
- [40] D. J. Ewins. *Modal testing: theory, practice, and application*. Research Studies Press, 2000.
- [41] Carl F. Eyring. "Reverberation Time in "Dead" Rooms". In: *The Journal of the Acoustical Society of America* 1.2A (1930), pp. 168–168. DOI: 10.1121/1.1901884. eprint: <https://doi.org/10.1121/1.1901884>. URL: <https://doi.org/10.1121/1.1901884>.
- [42] Bruno Fazenda. "Perception of Room Modes in Critical Listening Spaces". PhD thesis. 2004.
- [43] George E. Forsythe. "Generation and Use of Orthogonal Polynomials for Data-Fitting with a Digital Computer". In: *Journal of the Society for Industrial and Applied Mathematics* 5.2 (1957), pp. 74–88. ISSN: 03684245. URL: <http://www.jstor.org/stable/2098723>.
- [44] W. Frommhold, H.V. Fuchs, and S. Sheng. "Acoustic Performance of Membrane Absorbers". In: *Journal of Sound and Vibration* 170.5 (1994), pp. 621–636. ISSN: 0022-460X. DOI: <https://doi.org/10.1006/jsvi.1994.1091>. URL: <https://www.sciencedirect.com/science/article/pii/S0022460X84710911>.
- [45] Mallat Stéphane G. and Peyré Gabriel. *A wavelet tour of signal processing: the sparse way*. Elsevier, 2009.
- [46] Patrick R. Gill, Albert Wang, and Alyosha Molnar. "The In-Crowd Algorithm for Fast Basis Pursuit Denoising". In: *IEEE Transactions on Signal Processing* 59.10 (2011), pp. 4595–4605. DOI: 10.1109/TSP.2011.2161292.
- [47] Claude Gittelsohn, Ralf Hiptmair, and Ilaria Perugia. "Plane wave discontinuous Galerkin methods". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 43 (Mar. 2009), pp. 297–331. DOI: 10.1051/m2an/2009002.
- [48] G.m.l. Gladwell and N.b. Willms. "On The Mode Shapes Of The Helmholtz Equation". In: *Journal of Sound and Vibration* 188.3 (1995), pp. 419–433. DOI: 10.1006/jsvi.1995.0602.
- [49] Michael Goodwin and Martin Vetterli. "Matching pursuit and atomic signal models based on recursive filter banks". In: *Signal Processing, IEEE Transactions on* 47 (Aug. 1999), pp. 1890–1902. DOI: 10.1109/78.771038.
- [50] B. Gustavsen and A. Semlyen. "Rational approximation of frequency domain responses by vector fitting". In: *IEEE Transactions on Power Delivery* 14.3 (1999), pp. 1052–1061. DOI: 10.1109/61.772353.
- [51] Bjørn Gustavsen and Christoph Heitz. "Modal Vector Fitting: A Tool For Generating Rational Models of High Accuracy With Arbitrary Terminal Conditions". In: *Advanced Packaging, IEEE Transactions on* 31 (Dec. 2008), pp. 664–672. DOI: 10.1109/TADVP.2008.927810.

BIBLIOGRAPHY

- [52] Mark H. Richardson and David L. Formenti. “Global Curve Fitting of Frequency Response Measurements using the Rational Fraction Polynomial Method”. In: *International Modal Analysis Conference and Exhibit* (1985).
- [53] Dominique Habault et al. “Active Control in an Anechoic Room: Theory and First Simulations”. In: *Acta Acustica united with Acustica* 103 (May 2017), pp. 369–378. DOI: 10.3813/AAA.919066.
- [54] Y. Haneda, S. Makino, and Y. Kaneda. “Multiple-point equalization of room transfer functions by using common acoustical poles”. In: *IEEE Transactions on Speech and Audio Processing* 5.4 (1997), pp. 325–333. DOI: 10.1109/89.593306.
- [55] Ben Hanson and Martin Levesley. “Self-sensing applications for electromagnetic actuators”. In: *Sensors and Actuators A: Physical* 116.2 (2004), pp. 345–351. ISSN: 0924-4247. DOI: <https://doi.org/10.1016/j.sna.2004.05.003>. URL: <https://www.sciencedirect.com/science/article/pii/S0924424704003206>.
- [56] R. Hiptmair, A. Moiola, and I. Perugia. “Plane Wave Discontinuous Galerkin Methods for the 2D Helmholtz Equation: Analysis of the p-Version”. In: *SIAM Journal on Numerical Analysis* 49.1 (2011), pp. 264–284. DOI: 10.1137/090761057. eprint: <https://doi.org/10.1137/090761057>. URL: <https://doi.org/10.1137/090761057>.
- [57] Piotr Indyk. “Sparse Recovery Using Sparse Random Matrices”. In: Apr. 2010, p. 157. ISBN: 978-3-642-12199-9. DOI: 10.1007/978-3-642-12200-2_15.
- [58] Anna Izewska. “Measurement uncertainty of the sound absorption coefficient”. In: *The Journal of the Acoustical Society of America* 123.5 (2008), pp. 3501–3501. DOI: 10.1121/1.2934375. eprint: <https://doi.org/10.1121/1.2934375>. URL: <https://doi.org/10.1121/1.2934375>.
- [59] Lee K. Jones. “On a Conjecture of Huber Concerning the Convergence of Projection Pursuit Regression”. In: *The Annals of Statistics* 15.2 (1987), pp. 880–882. DOI: 10.1214/aos/1176350382. URL: <https://doi.org/10.1214/aos/1176350382>.
- [60] Sami Karkar et al. “Electroacoustic absorbers for the low-frequency modal equalization of a room: what is the optimal target impedance for maximum modal damping, depending on the total area of absorbers?” In: *Proceedings of Forum Acusticum* (2014).
- [61] Mendel Kleiner and Jiri Tichy. *Acoustics of small rooms*. CRC Press, 2014.
- [62] Bjørn Kolbrek and U. Peter Svensson. “Modeling Non-Shoebox Shaped Rooms with the Mode Matching Method”. In: *Journal of The Audio Engineering Society* (2016).
- [63] Shoichi Koyama and Hiroshi Saruwatari. “Sound field decomposition in reverberant environment using sparse and low-rank signal models”. In: *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2016, pp. 395–399. DOI: 10.1109/ICASSP.2016.7471704.
- [64] M. A. Kuczmarski and J.C. Johnston. “Acoustic Absorption in Porous Materials”. In: *NASA* (2011).
- [65] Heinrich Kuttruff. *Room acoustics*. CRC press, 2009.

-
- [66] Jenny Lau et al. “Automatic modal analysis. Reality or myth?” In: *Sound & vibration* (Jan. 2007).
 - [67] Q. Leclère, N.B. Roozen, and Céline Sandier. “On the use of the Hs estimator for the experimental assessment of transmissibility matrices”. In: *Mechanical Systems and Signal Processing* 43 (Feb. 2014), pp. 237–245. DOI: 10.1016/j.ymssp.2013.09.008.
 - [68] Yin Tat Lee and Aaron Sidford. “Efficient Inverse Maintenance and Faster Algorithms for Linear Programming”. In: *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*. 2015, pp. 230–249. DOI: 10.1109/FOCS.2015.23.
 - [69] Paul Leopardi. “A partition of the unit sphere into regions of equal area and small diameter”. In: *Electron. Trans. Numer. Anal.* 25 (Jan. 2006).
 - [70] Junhong Lin and Song Li. “Nonuniform support recovery from noisy random measurements by Orthogonal Matching Pursuit”. In: *Journal of Approximation Theory* 165.1 (2013), pp. 20–40. ISSN: 0021-9045. DOI: <https://doi.org/10.1016/j.jat.2012.09.009>. URL: <https://www.sciencedirect.com/science/article/pii/S0021904512001724>.
 - [71] Hervé Lissek, Romain Boulandet, and Romain Fleury. “Electroacoustic absorbers: Bridging a gap between active sound absorption and shunt loudspeakers”. In: *The Journal of the Acoustical Society of America* 129 (May 2011), pp. 2968–78. DOI: 10.1121/1.3569707.
 - [72] Hervé Lissek, Romain Boulandet, and Etienne Rivet. “Optimization of electric shunt resonant circuits for electroacoustic absorbers”. In: Apr. 2012. ISBN: 9782919340019.
 - [73] Niranjana Londhe, Mohan D. Rao, and Jason R. Blough. “Application of the ISO 13472-1 in situ technique for measuring the acoustic absorption coefficient of grass and artificial turf surfaces”. In: *Applied Acoustics* 70.1 (2009), pp. 129–141. ISSN: 0003-682X. DOI: <https://doi.org/10.1016/j.apacoust.2007.12.011>. URL: <https://www.sciencedirect.com/science/article/pii/S0003682X07002009>.
 - [74] Aki Mäkitvirta et al. “Low-frequency modal equalization of loudspeaker-room responses”. In: *AES: Journal of the Audio Engineering Society* 51 (Jan. 2001).
 - [75] S.G. Mallat and Zhifeng Zhang. “Matching pursuits with time-frequency dictionaries”. In: *IEEE Transactions on Signal Processing* 41.12 (1993), pp. 3397–3415. DOI: 10.1109/78.258082.
 - [76] Karjalainen Matti and Tuomas Paatero. “Equalization of Loudspeaker and Room Responses Using Kautz Filters: Direct Least Squares Design”. In: *EURASIP Journal on Advances in Signal Processing* 2007 (Jan. 2007). DOI: 10.1155/2007/60949.
 - [77] Manuel Melon et al. “Evaluation of a method for the measurement of subwoofers in usual rooms”. In: *The Journal of the Acoustical Society of America* 127.1 (2010), pp. 256–263. DOI: 10.1121/1.3270392. eprint: <https://doi.org/10.1121/1.3270392>. URL: <https://doi.org/10.1121/1.3270392>.
 - [78] Yves Meyer and D. H. Salinger. *Wavelets and operators*. Cambridge University Press, 2004.

BIBLIOGRAPHY

- [79] R. Mignot, G. Chardon, and L. Daudet. “Low Frequency Interpolation of Room Impulse Responses Using Compressed Sensing”. In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22.1 (Jan. 2014), pp. 205–216. ISSN: 2329-9290. DOI: 10.1109/TASLP.2013.2286922.
- [80] M. Mironov, Alexander Komkin, and Aleksei Bykov. “Sound Absorption by a Helmholtz Resonator”. In: *Acoustical Physics* 63 (July 2017), pp. 385–392. DOI: 10.1134/S1063771017030071.
- [81] M. Miyoshi and Y. Kaneda. “Inverse filtering of room acoustics”. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 36.2 (1988), pp. 145–152. DOI: 10.1109/29.1509.
- [82] A. Moiola, R. Hiptmair, and I. Perugia. “Plane wave approximation of homogeneous Helmholtz solutions”. In: *Zeitschrift für angewandte Mathematik und Physik* 62.5 (2011), pp. 809–837. DOI: 10.1007/s00033-011-0147-y.
- [83] Philip M. Morse and Richard H. Bolt. “Sound Waves in Rooms”. In: *Rev. Mod. Phys.* 16 (2 Apr. 1944), pp. 69–150. DOI: 10.1103/RevModPhys.16.69. URL: <https://link.aps.org/doi/10.1103/RevModPhys.16.69>.
- [84] Philip M. Morse and K. Uno Ingard. *Theoretical acoustics*. Princeton University Press, 1986.
- [85] Swen Müller and Paulo Massarani. “Transfer-Function Measurement with Sweeps”. In: *Journal of the Audio Engineering Society* 49 (June 2001).
- [86] Deanna Needell and Joel Tropp. “CoSaMP: Iterative Signal Recovery from Incomplete and Inaccurate Samples”. In: *Communications of the ACM* 53 (Dec. 2010). DOI: 10.1145/1859204.1859229.
- [87] R. Neff and A. Zakhor. “Very low bit-rate video coding based on matching pursuits”. In: *IEEE Transactions on Circuits and Systems for Video Technology* 7.1 (1997), pp. 158–171. DOI: 10.1109/76.554427.
- [88] oscar j. bonello oscar j. “a new criterion for the distribution of normal room modes”. In: *journal of the audio engineering society* 29.9 (Sept. 1981), pp. 597–606.
- [89] Y.C. Pati, R. Rezaeiifar, and P.S. Krishnaprasad. “Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition”. In: *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*. 1993, 40–44 vol.1. DOI: 10.1109/ACSSC.1993.342465.
- [90] Lueg Paul. “Process of silencing sound oscillations”. In: 2043416 (June 1936). URL: <https://www.freepatentsonline.com/2043416.html>.
- [91] Bart Peeters et al. “The PolyMAX frequency-domain method: A new standard for modal parameter estimation? Shock and Vibration”. In: *Special Issue Dedicated to Professor Bruno Piombo* 395 (Jan. 2004).
- [92] Laurent Perrinet, Manuel Samuelides, and Simon Thorpe. “Sparse spike coding in an asynchronous feed-forward multi-layer neural network using Matching Pursuit.” In: *Neurocomputing* 57 (July 2003), pp. 125–134. DOI: 10.1016/j.neucom.2004.01.010.

-
- [93] Daniel P. Petersen and David Middleton. "Sampling and reconstruction of wave-number-limited functions in N-dimensional euclidean spaces". In: *Information and Control* 5.4 (1962), pp. 279–323. ISSN: 0019-9958. DOI: [https://doi.org/10.1016/S0019-9958\(62\)90633-2](https://doi.org/10.1016/S0019-9958(62)90633-2). URL: <https://www.sciencedirect.com/science/article/pii/S0019995862906332>.
 - [94] Yaniv Plan and Roman Vershynin. "One-Bit Compressed Sensing by Linear Programming". In: *Communications on Pure and Applied Mathematics* 66 (Aug. 2013). DOI: 10.1002/cpa.21442.
 - [95] poju antsalo poju et al. "perception of temporal decay of low-frequency room modes". In: *journal of the audio engineering society* (May 2004).
 - [96] J. Ramis et al. "The Uncertainty in Absorption Coefficients Measured in Reverberant Chambers: A Case Study". In: *Noise & Vibration Worldwide* 36.1 (2005), pp. 7–12. DOI: 10.1260/0957456053499185. eprint: <https://doi.org/10.1260/0957456053499185>. URL: <https://doi.org/10.1260/0957456053499185>.
 - [97] Holger Rauhut. "Compressive Sensing and Structured Random Matrices". In: *Theor Found Numer Methods Sparse Recover* 9 (Jan. 2010). DOI: 10.1515/9783110226157.1.
 - [98] Mark H. Richardson and David L. Formenti. "Parameter estimation from frequency response measurements using rational fraction polynomials". In: *In Proceedings of the 1 st International Modal Analysis Conference*. 1982, pp. 167–181.
 - [99] Etienne Rivet. "Room Modal Equalisation with Electroacoustic Absorbers". PhD thesis. 2016.
 - [100] Etienne Rivet, Sami Karkar, and Herve Lissek. "Broadband Low-Frequency Electroacoustic Absorbers Through Hybrid Sensor-/Shunt-Based Impedance Control". In: *IEEE Transactions on Control Systems Technology* 25.1 (2017), pp. 63–72. DOI: 10.1109/tcst.2016.2547981.
 - [101] Etienne Rivet, Sami Karkar, and Herve Lissek. "On the Optimisation of Multi-Degree-of-Freedom Acoustic Impedances of Low-Frequency Electroacoustic Absorbers for Room Modal Equalisation". In: *Acta Acustica united with Acustica* 103.6 (2017), pp. 1025–1036. ISSN: 1610-1928. DOI: [doi:10.3813/AAA.919132](https://doi.org/10.3813/AAA.919132). URL: <https://www.ingentaconnect.com/content/dav/aaua/2017/00000103/00000006/art00016>.
 - [102] Etienne Rivet et al. "Experimental assessment of low-frequency electroacoustic absorbers for room modal equalization in actual listening rooms". In: *Audio Eng. Soc. Conv.*, 140 (June 2016).
 - [103] Olivier Robin et al. "Measurement of the absorption coefficient of sound absorbing materials under a synthesized diffuse acoustic field". In: *The Journal of the Acoustical Society of America* 136.1 (2014), EL13–EL19. DOI: 10.1121/1.4881321. eprint: <https://doi.org/10.1121/1.4881321>. URL: <https://doi.org/10.1121/1.4881321>.

BIBLIOGRAPHY

- [104] Paolo Romano and Mario Paolone. “Enhanced Interpolated-DFT for Synchronphasor Estimation in FPGAs: Theory, Implementation, and Validation of a PMU Prototype”. In: *IEEE Transactions on Instrumentation and Measurement* 63.12 (2014), pp. 2824–2836. DOI: 10.1109/TIM.2014.2321463.
- [105] W C Sabine. *Architectural Acoustics*. (1900; reprinted by Dover, New York, 1964).
- [106] Manfred R. Schroeder. “The “Schroeder frequency” revisited”. In: *The Journal of the Acoustical Society of America* 99.5 (1996), pp. 3240–3241. DOI: 10.1121/1.414868. eprint: <https://doi.org/10.1121/1.414868>. URL: <https://doi.org/10.1121/1.414868>.
- [107] A. Skodras, C. Christopoulos, and T. Ebrahimi. “The JPEG 2000 still image compression standard”. In: *IEEE Signal Processing Magazine* 18.5 (2001), pp. 36–58. DOI: 10.1109/79.952804.
- [108] U. Peter Svensson. “Finding the resonance frequencies of non-shoebox shaped rooms”. In: 2019.
- [109] Floyd E. Toole. *Sound reproduction: the acoustics and psychoacoustics of loudspeakers and rooms*. Routledge, 2018.
- [110] Ivana Tomic, Pascal Frossard, and Pierre Vandergheynst. “Progressive Coding of 3-D Objects Based on Overcomplete Decompositions”. In: *IEEE Transactions on Circuits and Systems for Video Technology* 16.11 (2006), pp. 1338–1349. DOI: 10.1109/TCSVT.2006.883502.
- [111] Joel A. Tropp, Anna C. Gilbert, and Martin J. Strauss. “Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit”. In: *Signal Processing* 86.3 (2006), pp. 572–588. DOI: 10.1016/j.sigpro.2005.05.030.
- [112] M. Unser and T. Blu. “Mathematical properties of the JPEG2000 wavelet filters”. In: *IEEE Transactions on Image Processing* 12.9 (2003), pp. 1080–1090. DOI: 10.1109/TIP.2003.812329.
- [113] Jian Wang, Seokbeop Kwon, and Byonghyo Shim. “Generalized Orthogonal Matching Pursuit”. In: *IEEE Transactions on Signal Processing* 60.12 (2012), pp. 6202–6216. DOI: 10.1109/TSP.2012.2218810.
- [114] Yanghua Wang. “Multichannel matching pursuit for seismic trace decomposition”. In: *Geophysics* 75 (July 2010), pp. V61–V66. DOI: 10.1190/1.3462015.
- [115] Zhifeng Zhang. “Matching pursuit”. PhD thesis. 1993.
- [116] J. Zhu. “Image compression using wavelets and JPEG2000: A tutorial”. In: *Electronics Communication Engineering Journal* 14 (July 2002), pp. 112–121. DOI: 10.1049/ecej:20020303.

Curriculum Vitae

Thach Pham Vu was born in Hanoi, Vietnam. He received his BSc degree in Mechanical Engineering at Korea Advanced Institute of Science and Technology (KAIST) in Deajeon, South Korea in 2014. He was then admitted to École polytechnique fédérale de Lausanne (EPFL), Switzerland in 2014 for the Master program and later received his MSc in Mechanical Engineering in 2016. In November 2016, he enrolled in the PhD program at the Acoustics Group of the Signal Processing Laboratory (LTS2) at EPFL. As a teaching assistant, he has taken part in multiple teaching activities for classes on Audio Engineering and Electroacoustics. He also took part in the supervision of multiple student projects at the Acoustics Group of LTS2. He is a member of the French Acoustical Society.

List of publications

Journal papers

T. Pham Vu and H. Lissek, Low frequency sound field reconstruction in a non-rectangular room using a small number of microphones, *Acta Acust.* 4 (2) 5 (2020) DOI: 10.1051/aacus/2020006

Conference papers

T. Pham Vu and H. Lissek, Sound field reconstruction in a room using low-rank approximation of mode shapes: validation and application, *Proceedings of the 26th International Congress on Sound and Vibration (ICSV26)*, 2019

T. Pham Vu, E. Rivet and H. Lissek, Low Frequency Sound Field Reconstruction in Non-rectangular Room: A Numerical Validation, *Proceedings of Euronoise 2018, Crete, Greece 2018*.

H. P. Tukuljac, T. Pham Vu, H. Lissek and P. Vanderghenst, "Joint Estimation of the Room Geometry and Modes with Compressed Sensing," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018, pp. 6882-6886, doi: 10.1109/ICASSP.2018.8462655.

T. Pham Vu, E. Rivet and H. Lissek, Sound Field Reconstruction In Low-Frequency Room Acoustics: A Benchmark Study With Simulation, *proceedings of COMSOL Conference, Lausanne, 2018*