UNIVERSITY OF SOUTHAMPTON

Faculty of Engineering and Physical Sciences
Institute of Sound and Vibration Research

Signal processing strategies to improve low-frequency performance of multi-channel systems in cuboid reverberant spaces.

by

Thomas Oliver Bell

Thesis for the degree of Doctor of Philosophy

November 2023

University of Southampton

Abstract

Faculty of Engineering and Physical Sciences
Institute of Sound and Vibration Research

Thesis for the degree of Doctor of Philosophy

Signal processing strategies to improve low-frequency performance of multi-channel systems in cuboid reverberant spaces.

Thomas Oliver Bell

This thesis addresses problems in the reproduction of low-frequency audio, in home listening environments, using digital signal processing. Focusing on cuboid rooms, it covers two primary areas. Firstly, controlling the sound field within an enclosure through a least-squares approach to filter design, secondly, predicting this field from a minimal number of acoustic measurements. It concludes that filters designed using a four-channel least-squares approach, based on pseudo-real-world data, and created using the proposed enhanced analytical model, show a measurable improvement in the real world.

A literature review is presented of the physical properties of sound within an enclosure, various previous approaches addressing low-frequency issues, and methods for understanding and predicting the sound field. From this review, it is clear that the proposed approach for predicting the sound field within an enclosure is unique. The plane wave cancellation method offers a good solution but with limitations, and this is very thoroughly examined. The target is clearly to create a system objectively as effective as the plane wave cancellation method but without its limitations. The least-squares filter approach is introduced, and shown through numerical simulations to outperform the plane wave cancellation method, without its limitations. It is also shown that the effectiveness of the least-squares filter approach seen in numerical simulations, transfers to real-world data.

Finally, methods are proposed for extracting room dimensions and source and receiver locations, from the frequency analysis of a small number of real-world acoustic measurements. The results are combined with Green's function, to create an enhanced analytical model, that can give pseudo-real-world measurements, accurate enough to replace the need for real-world measurements, when designing least-square filters.

Contents

List of Figures

List of Tables

Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:
    - T. O. Bell and F. M. Fazi. An Investigation into the Location and Number of Microphone Measurements Necessary for Efficient Active Control of Low-Frequency Sound Fields in Listening Rooms. In *Audio Engineering Society Convention 147*, New York, USA, October 2019. AES, 132 East 43rd Street, Suite 405, NY, NY 10017
    - T. O. Bell and M. Baratelli. A Comparison Between Plane Wave Cancellation and a Least-Squares Approach for the Zonal Control of Modes in a Rectangular Room. In *Audio Engineering Society Convention 148*, Online, June 2020. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=20759

Signed:..... Date:.....

Acknowledgements

The first person I would like to thank is my supervisor, Prof. Filippo Fazi, for enduring me for almost eight years in total. His immense knowledge and guidance have been irreplaceable. He has always been nothing but encouraging and positive. When we started, I felt lucky that someone with such a pedigree would even accept me as a student, and now we've finished, I feel lucky to have been able to work with you.

I would also like to thank my secondary supervisor, Prof. Steve Elliot, for your input and for taking over while Filippo was on sabbatical.

A massive thanks goes to Albert Yong for thinking the PhD was a good idea to start with. To Dr. Gary Geaves for agreeing and get this process started. To Dr. Martail Rousseau, who was forced to carry the can when both Gary and Albert left the UK, dumping the responsibility on him, your support and input were invaluable.

As a part-time remote-learning PhD student, it was always going to make it hard to get to know my fellow students, but I'd like to thank: Dr. Falk-Martin Hoffmann, Dr. Olivieri, Dr. Fabio Hirono, Dr. Marcos Gálvez, Dr. Eric Hamdan, and Dr. Dylan Menzies. You were always so welcoming, friendly and helpful on my infrequent campus visit. So infrequent was my presence in the group meeting for the first year Ferdinando was convinced Filippo had created an imaginary PhD student.

On a personal note, I'd like to thank my wife, Dr. Shabnah Ratnarajah (the actual medical type) for her unwavering love and support. You have been my rock and I hope one day to believe in myself the way you believe in me.

So much thanks to my Father Ian and Mother Jane who set me on this path and always wanted me to find my own way, never stopping supporting and encouraging me and to my mother sister Rosie. with thanks to Harvey the dog, and the hamsters Jack and Caroline.

To my friends: George Weaver, Michael Holman, Marco Baratelli Loe Lay, Dr. Paul Bryan and Cody Krieger I owe you all a debt for the nights of keeping me sane over a beer.

Dr. Mark Kermode and Simon Mayo, there have been times (before meeting my wife) that listening to your podcast on my long weekend runs has been the only light in constant weeks of work and study, 'Hello to Jason Isaacs'.

For my wife Shabnah and my parents,
you have taught so much and asked for so little,
I will keep listening and keep learning.

"If you loved me 'til my eyes got no more shine for you
If you walked beside me all the long way home
If you wasted all of your time on me
Well, I couldn't love you more
Just couldn't love you more
I couldn't love you more."
— John Martyn

Definitions and Abbreviations

Latin letters

A	Infinite rigid boundary	
a	Target transfer functions	
B	Infinite rigid boundary	
C	Matrix of transfer functions	
c_0	Speed of sound	$343\,ms^{-1}$
d	Pressure target	
e	Pressure error	
f	Frequency	
h	Filter vector in frequency domain	
i	$\sqrt{-1}$	
j	Cost function	
j_{vec}	Cost function for the vector method of compensation	
j_{met}	Cost function for the matrix method of compensation	
k	Wave number	
L_p	Sound Pressure level	$[dB]$
L	Length	
$\mathbf{l} = [l_1, l_2, l_3]$	Vector of dimensions of an enclosure	
M	Maximum number of receivers	
m	Receiver number	
n	Frequency index	

$\mathbf{n} = [n_1, n_2, n_3]$ Mode number (n_1, n_2, n_3)

p Acoustic pressure

p_0 Scalar to set target level

$\tilde{p}_w$ Either the modelled or measured pressures at receiver points

Q Volume velocity

$\hat{Q}$ Q factor of individual modes

r Receiver

r Vector of co-ordinate positions for receiver location in the enclosure: (x, yz)

$\hat{\mathbf{r}}$ Vector of co-ordinate positions for receiver location built into source in the enclosure: $(\tilde{x}, \tilde{y}, \tilde{z})$

s Source

s Vector of co-ordinate positions for source location in the enclosure: $(\tilde{x}, \tilde{y}, \tilde{z})$

T $_n$ Duration of time (Chapter 6)

t Time

u Source output signal

V Volume

v Input signal

W Maximum number of sources

w Source number

$\hat{W}$ Maximum number of receivers built into sources

$\hat{w}$ Receiver built into source number

x, $\tilde{x}$ Cartesian coordinate from point in room along l_1 axis

$x = [x_1, x_2, x3]$ Vector of co-ordinate positions in the enclosure

y, $\tilde{y}$ Cartesian coordinate from point in room along l_2 axis

z, $\tilde{z}$ Cartesian coordinate from point in room along l_3 axis

Greek letters

β Tikhonov regularisation parameter

δ Dirac delta function

ε	Error	$[dB]$
$\bar{\varepsilon}$	Mean error	$[dB]$
ζ	Damping coefficient	
λ	Wavelength	
τ_n	The time at which the peak due to the harmonic distortion appears before the main impulse	
ϕ	Phase	
ρ_0	Air density	$1.225 kgm^{-3}$
ω	Angular frequency	
v	Filter vector in frequency domain for the vector method of compensation	
Y	Filter matrix in frequency domain for the matrix method of compensation	

Abbreviations

$2C_F$	Two-channel (front) approach
$2C_R$	Two-channel (rear) approach
3C	Three-channel approach
4C	Four-channel approach (Chapter 4)
4C	Results of filters designed with only real-world data (Chapter 5)
4C $_{Model}$	Four-channel approach results modelled with real-world data
4C $_{Real}$	Four-channel approach results with real-world data
4C	Results of filters designed with only real-world data (Chapter 5)
$4C_{RT60}$	Results of filters designed with mixed real-world and analytical model data
$4C_{4MT60+}$	Results of filters designed with mixed real-world and enhanced analytical model data
$4C_{5MT60+}$	Results of filters designed with only enhanced analytical model data
AES	Audio Engineering Society
C	Layout ideal for plane wave control method
CABS	Controlled active bass system
D	Layout with dual listening areas

EBU European Broadcasting Union

FA Average frequency response

ITU International Telecommunication Union

LT Long thin room

MD Magnitude deviation

MSFD Mean sound field deviation

MSV Mean spatial variance

MT_{60} Modal decay time

MT_{60+} Combination of the matrix method with modal decay time

O Layout where s_1 and s_2 are offset

PWC Plane wave cancellation method signal

RMS Root mean square

RS Raw signal

RT_{60} Reverberation time

S Layout where s_1 and s_2 are moved into a stereo placement

SA Spatial average

SD Spatial deviation

SPL Sound pressure level

SW Short wide room

T Layout with triple listening areas

WFS Wave field synthesis

Chapter 1

Introduction

This thesis addresses the problems in the reproduction of low-frequency audio in compromised home listening environments using digital signal processing. There are two main factors that affect low-frequency sound reproduction in enclosed spaces. The first is the modal resonances. These are frequencies at which the dimensions of an enclosed space create standing waves. At low-frequencies when the modal density is low, these resonances create isolated areas of frequency boost or cut, that can have a noticeable effect on the listening experience (Elliott (2000)). The second is reverberation, which is the time the acoustic energy takes to disperse to average background levels (Everest (1989)). Both factors are affected by the absorption characteristics and location of materials found within the space, and both factors control the clarity, perceived tonal balance, and the image definition of the audio. Methods for controlling the acoustic characteristics of environments through digital signal processing have been an area of ongoing research and development for over 40 years (Neely and Allen (1979)). However, as discussed in Cecchi et al. (2017), although numerous approaches are available, each approach has its advantages and limitations, and there remains no consensus on the ideal method. To further the understanding in this area, there are two overarching themes in this thesis.

1. Achieving the level of control seen when using the plane wave cancellation method as presented by Celestinos and Nielsen (2008b), while being less constrained by the source layout.

2. Extracting the key physical parameters of a cuboid listening environment from a small set of acoustic measurements, in order to build an enhanced analytical model that can predict transfer functions accurately enough to minimise the requirement for the end-user to take multiple real-world measurements.

The structure of this chapter is as follows:

- **Background:** Discussion into the wider motivations for this work. The points and assertions raised in this section are reviewed and developed with the appropriate evidence in Chapter 2.
- **Focus:** Defines the research question and sets out the objectives of this work.
- **Summary of the thesis:** An overview of the structure and flow of the thesis, and a summary of the information in each chapter.
- **Scientific contributions:** Outlines the specific unique scientific contributions of this thesis.

1.1 Background

In home audio reproduction, the effect the environment has on the audio creates an issue: the listener does not hear the audio as intended. Instead, the audio is coloured by the combination of the Hi-Fi equipment, the listening room's acoustic response and the relative placement of the loudspeakers and the listener. To achieve the perfect listening experience, large sums of money can be spent on equipment and acoustic treatments for listening rooms. This typically requires the room to be set aside for this single use only, putting the cost outside most people's budget. Generally, the listening environment is the main living room; it is normally multipurpose and, possibly, an open-plan space, so balance is required. Loudspeakers and listening positions are usually compromised for ease of living, and acoustic treatments lose out to more desirable decor. The main aim of a room control system is to give the listener the closest possible experience to the perfect room in this compromised environment. To achieve this, the majority of room correction approaches need to understand the environment for which they are aiming to compensate. This is normally done with physical measurements taken in the environment (Bharitkar and Kyriakakis (2011); Bharitkar (2003); Pedersen (2012)), leading to one of the central issues of these control systems, namely that they require some action of setup by the end-user. Hence, the effectiveness of the control is reliant on the amount of time and effort the end-user is willing to put in, and once complete, the setup and areas of control are fixed, unless the end-user is willing to rerun the process. Therefore, the ability to accurately predict physical measurements within the environment with a mathematical model will lessen the end-user effort needed and could mean that the listening area(s) can be moved around the room with little or no physical measurement.

In the last five years, there has been a rise in smart systems for the home. This has led to interconnected audio systems in which each individual loudspeaker has its own on-board microphone and audio digital signal processor (DSP). These smart loudspeakers can work both as individual nodes with unique processing monitoring, modifying their

own output in real-time, and as part of the networked cluster to understand the system as a whole and carry out group processing. This change in each loudspeaker's hardware coupled with the year-on-year improvements in microprocessor technology, best outlined by Moore's law (Moore (1965)), means that the available processing power has increased significantly. Not only can the onboard DSP run real-time sample-by-sample algorithms, but the whole system can act as a cluster computer wherein individual systems are grouped in a network to work in parallel (Amdahl (1967)). It can offer more mathematical support for less time-critical processes and even allow the system to pass data to the cloud for heavy mathematical processes for which time is not a factor. A non-time-critical process is one that does not need to be run at the same clock rate as the DSP core; for example, monitoring changes in the room acoustics over time and making subsequent filter adjustments. The long-term analysis and decision-making should not affect audio playback and therefore should not be handled by the time-critical audio path in the DSP. Once a decision is made to adjust the filter and new coefficients are calculated, the filters in the time-critical audio path are updated.

With this unprecedented amount of shared processing power, it is time to review what the current room control solutions offer, assess whether there is a better method and understand what extra processing can do to improve models based on real-world data to predict in-room sound fields and identify whether they can even remove the need for listener-reliant setup processes.

1.2 Focus

The wider motivation outlined in the previous section leads to the research question, which can be framed as:

> Can a multichannel audio system be created that understands and predicts the sound field in a listening environment from a small set of acoustic measurement points, and then compensates for issues produced over the listening area due to non-ideal loudspeaker placements and environmental acoustic anomalies, balancing consumers' want for the highest quality sound reproduction, while offering minimal impact on the living space?

To answer this question, it is first necessary to split it into two parts. Then, for each part, there is a defined set of objectives.

Part 1 : Compensating for acoustic issues over the listening area in an enclosure.

The objectives of this part are:

(a) To review the acoustic issues faced within listening environments, especially at low frequencies, and to outline the previous methods that compensated for these acoustic issues and define a best-in-class reference method of control from these previous methods.

(b) To gain a deeper understanding of the reference method of control and show that there is a limitation in the reference method that a new approach can address.

(c) To propose a new approach and show it to be at least as effective as the reference method, but without the same limitations.

Part 2 : Understanding and predicting the sound field in a cuboid enclosure. In **Part 1** full knowledge of the listening environment has been assumed, in the real world this may not be the case. Therefore, the focus of this part is how the system can gain the required understanding of the acoustic properties of the environment in order to compensate for acoustic issues.

The objectives for this part are:

(a) To review the previous methods to understand and predict the sound field in the listening environment.

(b) To understand if the physical properties of the enclosed listening environment (including the dimensions, loudspeaker placements and receiver location) can be extracted through the analysis of acoustic measurements.

(c) To investigate whether an enhanced analytical model can accurately predict the measured transfer functions once all the relevant physical properties of the real-world enclosed listening environment are known.

Combining these two parts will give an effective method for controlling acoustic issues faced within cuboid listening environments, with the enhanced analytical model replacing the need for multiple acoustic measurement points across the listening area.

1.3 Thesis Structure

This section provides an overview of the structure of the thesis and briefly summarises the work in each chapter.

Chapter 2

This chapter provides the reader with an understanding of the subjects and tools required to continue through the thesis. There are four main sections in this chapter:

1. A review of room acoustics to allow full understanding of the modal issues found in enclosed spaces and how they can be modelled and predicted.
2. A review of previous work in the field of room acoustic control, with a focus on the plane wave cancellation method.
3. The introduction of metrics for the analysis of acoustic control.
4. A review of previous work in the area of understanding and predicting sound fields in listening environments.

Chapter 3

Having introduced the method of plane wave cancellation in Chapter 2, this chapter studies the method in more detail and consists of

1. An in-depth review of the theories underpinning the plane wave cancellation method.
2. Results from studies into the effectiveness of the plane wave cancellation method.
3. Results from studies that show how the plane wave cancellation method is constrained and how this may limit its usage in the home environment.

This chapter also introduces the analytical model used throughout the thesis.

Chapter 4

This chapter compares the plane wave cancellation method with four proposed approaches, each utilising the methods of pressure matching (Kirkeby and Nelson (1993)) and least-mean-squares filter design (Elliott and Nelson (1989)), with the aim of developing an approach that offers a comparable level of control with a less rigid system setup. The underlying mathematics in the four proposed approaches are not new; however, the specific usage is novel. This chapter is split into four sections, and the areas covered by each section are:

1. **Proposed method** – Considering the limitations of the plane wave cancellation method shown in Chapter 3, four least-squares approaches to filter design are proposed, each with varying degrees of freedom. The effectiveness of each approach is compared with the plane wave cancellation method using the analytical model.

2. **Real-world testing of the proposed method** – An investigation into how the most effective of the least-squares approaches transfers to the real world is conducted.

3. **Changing the source locations** – As shown in this thesis, the plane wave cancellation method is constrained by the source location. This section reports on a set of studies carried out as part of this work. These studies compared the plane wave cancellation method with the three most effective least-squares approaches over a range of source locations to understand if each method has the same constraint.

4. **Least-squares approach limitations** – A study to understand how the most effective of the least-squares approaches is affected by a range of situations is conducted in order to understand what limits it may have in controlling the sound field in a cuboid listening environment and how this may affect its possible usage in the home environment.

Chapter 5

This chapter focuses on understanding and predicting the sound field in a cuboid listening environment. As stated, the aim is to create an accurate analytical model of a cuboid enclosure to be able to predict the frequency response at any location.

The analytical model is the Green's function of the Helmholtz equation. In order to calculate transfer functions between two points in a cuboid enclosure, the Green's function must include coefficients for the dimensions of the enclosure, the locations of the source and receiver, and the damping of the surface of the enclosure. The theory is that if all of these coefficients are known, the Green's function will give an accurate analytical model of the real-world transfer function in the cuboid enclosure under test.

This chapter is split into two main parts. The first aims to estimate or compensate for all the coefficients from the analysis of the acoustic measurement data. They are:

1. **Damping** – During the testing of the Green's function, it became clear that the damping coefficient was the highest priority. If the damping coefficient was wrong, knowing all the other coefficients would not provide a sufficiently accurate result. To address this, five methods are compared in this section.

    - Three methods estimating the damping coefficient within the Green's function.
    - Two methods that do not estimate the damping coefficient within the Green's function, but try to compensate for the unknown damping coefficient by filtering the Green's function with filters derived from real-world measurements.

This work is carried out with known dimensions and locations to prove the theory. However, in the final algorithm, the damping estimation comes after the dimensions of the room and the locations of the sources and receivers have been estimated.

2. **Dimensions** – Introducing and testing the algorithmic process that estimates the dimensions of enclosures from frequency response data.

3. **Locations** – Introducing and testing the algorithmic process using constrained least-squares and trigonometric calculations to estimate the locations (in Cartesian coordinates) from the frequency response data. This process used the previously estimated dimensions.

The second part combines the work from the start of this chapter with the work from Chapter 4 to test how effective the four-channel least-squares approach is when only one real-world measurement point is in the listening area. This single measurement point is combined with four measurement points for which data has been created using the enhanced analytical model and then tested with the same process outlined in Chapters 3 and 4.

Chapter 6

The final chapter presents a summary of the conclusions drawn from the work presented in the preceding chapters.

1.4 Contributions

The original contributions of this thesis are:

1. An in-depth comparison of the plane wave cancellation method and the least-squares approach in the control of room acoustics. A rigorous study of variants of the latter through numerical simulations and real-world experiments.

2. A method for determining room dimensions, audio source and receiver locations in a cuboid environment from the frequency analysis of real-world measurements. Furthermore, a comparison of methods of estimating damping coefficients for use with the Green's function and the introduction of two methods of damping compensation.

3. The derivation of an analytical model that outputs data close enough to real-world measurements to provide effective control of room acoustics when using

the four-channel least-squares approach, and hence removes the need for any real-world measurements within the listening area.

Chapter 2

Previous work

This chapter aims to outline the background literature and current understanding in the research undertaken in the later chapters of this thesis.

There are Four main sections.

1. The relevant research on room acoustics.
2. Current active systems that aim to offer a level of acoustic control.
3. The introduction of metrics for the analysis of acoustic control.
4. Current systems for measuring and understanding room environments.

2.1 Room acoustics primer

It was stated in the introduction that audio is coloured by the combination of the hi-fi equipment and loudspeakers used and the acoustic response of the listening room. One main area of interest is how the listening environment affects the audio experience. This section will consider the effects of the room and why it can create problems.

2.1.1 Sound propagation

Sound in a free field propagates equally in all directions, with the level of the propagating sound being attenuated over distance due to energy spreading and the absorption of air (Beranek and Mellow (2019)). However, once in an enclosure, the boundaries of the enclosure and any objects within it affect the propagation path. How the sound is affected depends on the frequency of the signal and the surface type, size, shape and material of any objects or boundaries (Everest (1989)).

Throughout this thesis, focus is given only to the effect the boundaries of the enclosure have on the acoustic energy. This effect can be divided into two categories; either the energy is reflected into the enclosure or not. The focus of this thesis is reflected energy. The term 'absorbed' is used as a catch-all term for lost energy, that is, the energy lost at the boundaries, be it through any method, including but not limited to, transmission, porous absorption, or resonant absorption (Everest (1989)). Additionally, if the boundary walls are considered rigid, all the energy will be reflected into the enclosure. A rigid boundary is a theoretical tool; in reality, there will always be a slight loss of energy, even if it is due only to the heat transfer between the boundary and the particles hitting it.

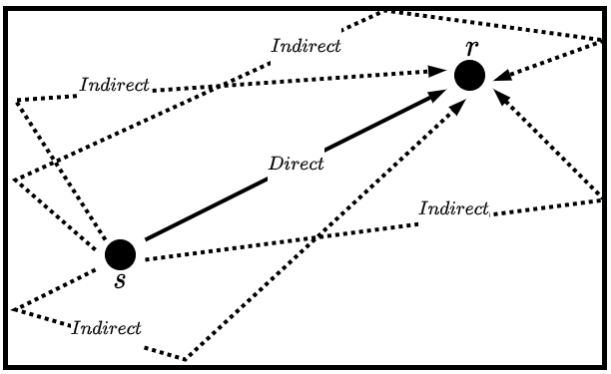
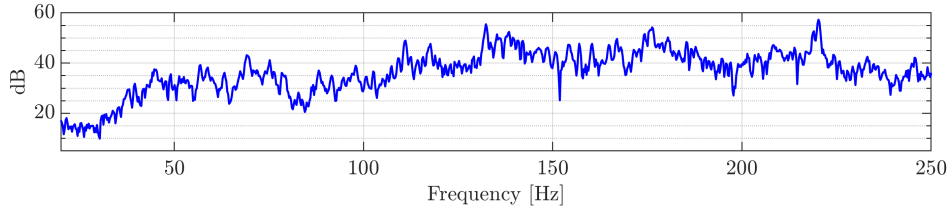


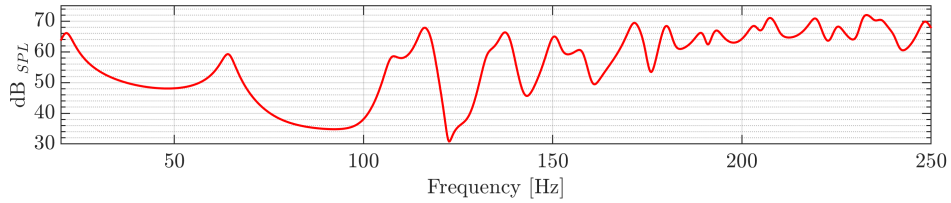
FIGURE 2.1: Propagation of sound within an enclosure, from a source at position **s** to a receiver at position **r**.

Figure 2.1 shows a representation of sound propagating in an enclosure from a source at position **s** to a receiver at position **r**, illustrating the direct sound path and a small number of indirect paths reflected off the boundaries. With non-rigid boundary walls, each time the wave is reflected from the surface, energy is lost, and the audio paths will continue throughout the enclosure until all the energy is absorbed. Additionally, depending on the nature of the boundary, the energy loss will be non-uniform across the frequency band, although boundary absorption normally increases with frequency in a linear fashion (Everest (1989)). As can be seen in Figure 2.1, these paths are of different lengths, meaning that a signal from **s** arrives at **r** over a time range depending on the length of the path travelled and the time taken for the energy to be absorbed. Due to the phase shift and loss of frequency from each boundary interaction, a listener at point **r** will hear differences in the arriving audio compared to the audio that left point **s** (Toole (1986)).

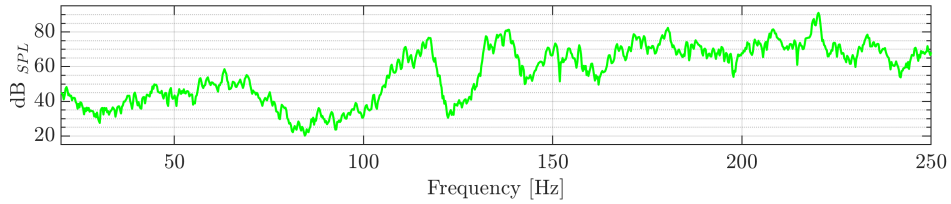
Figure 2.2 shows three basic magnitude plots to help explain how this affects the audio. Plot 2.2a shows the spectrum of an audio signal reproduced from **s**, which is what a listener would hear should the source be a pair of headphones or a loudspeaker in an anechoic room. Plot 2.2b shows the room's frequency response at the receiver position, and Plot 2.2c shows what a listener at the point *r* would actually hear (i.e. the combination of the audio and the room).



(a) Loudspeaker output signal at point **s** .



(b) Room response at point **r** .



(c) Audio response at point **r** due to signal from point s.

FIGURE 2.2: Room effect.

Plot 2.2c shows that the room acts as a filter for the audio. Each room offers a unique filter shape depending on the enclosure dimensions and the absorption of the walls. A metric to quantify the time taken for the energy to be absorbed, known as the reverberation time [RT_{60}], was defined by Sabine (1922) as the time that it takes an impulse response to decay 60 dB.

When creating audio for reproduction, engineers take the effects various rooms have into consideration (Jackson and Leventhall (1972); Burgess and Utley (1985)). The aim is to create a balanced audio mix that works well across a range of listening situations and environments, including but not limited to in the home, car, headphones and broadcast, all with both stereo and mono playback. To achieve this, engineers mix in the recording studio, and listen to mix-downs in other environments to obtain a balance (Alkin (1996) and Rumsey and McCormick (2002)). Clearly, not all possibilities can be tested or considered. The main target is to mix the audio so that it will sound best in a well-designed listening room, and to this end, recording studio mix-rooms are normally designed to well-accepted standards, such as those proposed by the European Broadcasting Union (EBU) in Hoeg et al. (1997) and AES (2001). Generally, the target RT_{60} is between 0.2 s to 0.5 s from 250 Hz to 4 kHz, with studies such as that by Beranek and Mellow (2019) stating that at frequencies below 250 Hz, this target may rise by 1.25%.

2.1.2 The frequency band of interest

The principal focus of this work is lower-frequency audio below 100 Hz. The boundary between what is classified as low- and high-frequency audio changes depending on the space under examination. Generally, it is the point where the acoustic response of the space moves from clearly individual resonances (or modes) to multiple overlapping resonances, which is known as the Schroeder frequency (Schroeder (1996)):

$$f_s = \sqrt{\frac{3c_0^3 T_r}{4\pi V \cdot 2.2}} \approx 2000\sqrt{\frac{T_r}{V}}. \tag{2.1}$$

In the above equation, V is the room's volume in m^3, and T_r is the reverberation time in seconds. Above the Schroeder frequency, the modal density and overlap are such that individual mode calculation is rarely useful. Instead, it is assumed that the pressure at any given point is the average of the superposition of waves travelling to that point from all possible directions (Nelson and Elliott (1992)). In the region below the Schroeder frequency, the individual modes are normally calculated. A good listening environment requires not only for the RT_{60} to be within the target values, but also a regular distribution of the modal frequencies in this low-frequency section of the audio spectrum (Louden (1971) and Meissner (2018)). This is demonstrated in Figure 2.2. As the frequency decreases, the mode peaks and troughs are defined in more detail in Plot 2.2b, and the effect this has on the audio signal is shown in Plot 2.2c.

2.1.3 Room modes

Room modes are well documented in many studies; this section was written based on Kinsler et al. (1982), Kuttruff (2017), Morse and Ingard (1968), Nelson and Elliott (1992), and Pierce (1981). A mode is a standing pressure wave between two or more surfaces, where the pressure is at maximum displacement from the mean at each surface. The result is a resonance at all frequencies with a wavelength equal to a multiple of twice the distance travelled. The fixed locations of these minimum and maximum pressures (relative to the standard average pressure within the medium) cause the sound at these frequencies to have different loudnesses at certain locations throughout a listening environment. In these standing waves, pressure and particle velocity are in quadrature, meaning areas of minimum and maximum pressure are areas of maximum and minimum particle velocity, respectively. The pressure level created in standing waves directly depends upon the absorption of the boundary at the resonance frequencies. If the boundary were to absorb all the energy at the modal frequencies, there would be no resonance; however, if it were to absorb no energy (assuming no other energy loss), the resonance would continue ad infinitum.

This section aims to build a mathematical framework to address the concept explained above. It begins in one dimension, which simplifies the mathematics, allowing the principles to be understood, and then moves into three dimensions. This mathematical review was conducted from basic first principles to obtain a fundamental understanding of the underlying process.

2.1.3.1 One-dimensional modes

A basic one-dimensional model provides a simple view of the mathematics behind modal resonances.

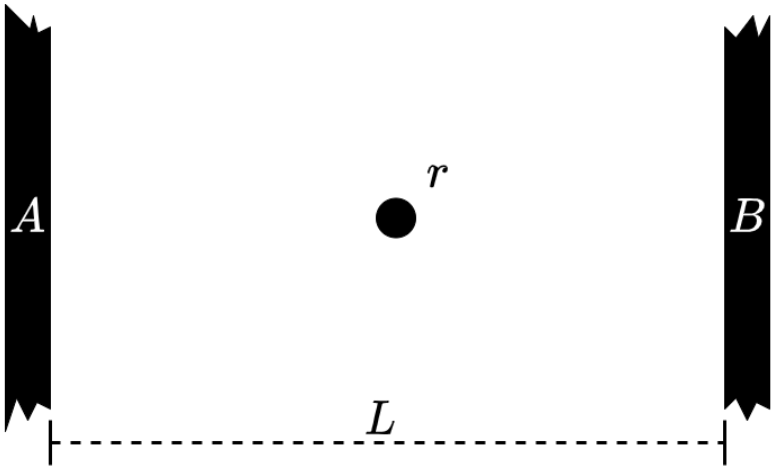


FIGURE 2.3: Basic one-dimensional model of acoustic pressure wave propagation between two surfaces.

Figure 2.3 shows a basic one-dimensional model for acoustic pressure wave propagation between two surfaces. A and B are separated by a distance L, and the following assumptions are made:

- A and B are infinite rigid boundaries.
- The medium between the boundaries allows for perfect transmission with no attenuation over distance.
- There is no heat loss in the system.
- The propagating wave is a complex pressure wave of infinite width with the wave-front form $p(x,t) = Qe^{i(\omega t - kx + \phi)}$. Here, Q is the volume velocity, ω is the angular frequency, t is the time, k is the wave number, x is the distance from the source, and ϕ is the phase.

If a single impulse (modelled as a Dirac delta function) is transmitted from surface A with the pressure measurement made at point r, where the distance from A to r is $\frac{L}{2}$, the

impulse should be reflected by the surfaces A and B continuously. If a timer is started such that $t = 0$ as the first impulse passes r, then r would measure a series of delta impulses over time.

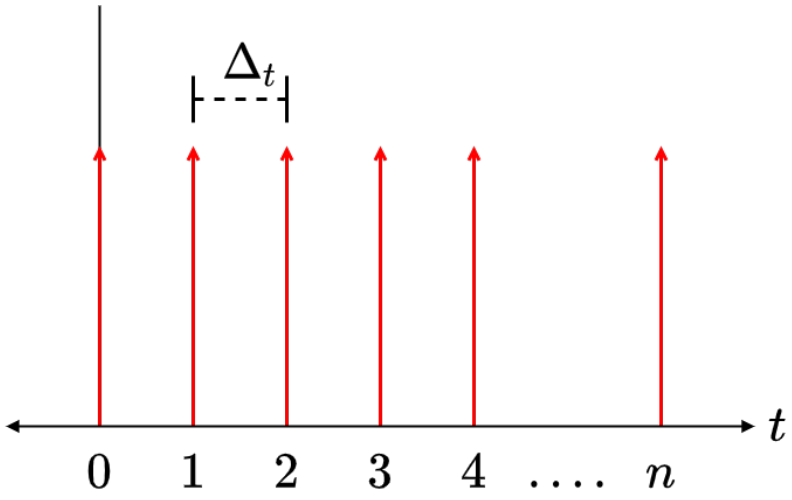


FIGURE 2.4: Train of delta impulses as measured at point p.

Figure 2.4 shows the measurement taken over time at point r. If $\Delta_t = \frac{L}{c_0}$, the sum of these of impulse can be expressed as:

$$p(t) = \sum_{n=0}^{N} \delta(t - n\Delta_t). \tag{2.2}$$

It was stated that the system has no loss. Therefore, once initiated, this impulse will continue for an infinite time. Rather than starting the measurement window at $t = 0$, it can be assumed that the system has been running since $t = -\infty$ and will run until $t = \infty$. The expression can then be rewritten as

$$p(t) = \sum_{-\infty}^{\infty} \delta(t - n\Delta_t), \tag{2.3}$$

the Fourier transform of which is

$$P(f) = \frac{1}{\Delta_t} \sum_{-\infty}^{\infty} e^{2\pi i n \Delta_t f}, \tag{2.4}$$

which, when evaluated with the Poisson sum formula (Williams (1999)) leads to

$$P(f) = \sum_{-\infty}^{\infty} \delta(f - \frac{n}{\Delta_t}). \tag{2.5}$$

This shows that a single impulse between two surfaces in the time domain converts into a train of impulses in the frequency domain. It also shows that the spacing between

the frequencies excited by this system depends inversely on the distance between the surfaces due to the fact that $\frac{n}{\Delta_t} = \frac{nc_0}{L}$.

Earlier in this section, it was stated that a mode is a standing pressure wave between two surfaces, where, at each surface, the pressure is at a maximum displacement from the mean. The pressure is

$$p(x,t) = \Re\{Qe^{i(\omega t - kx + \phi)}\}, \tag{2.6}$$

which, using Euler's formula, is

$$p(x,t) = Q\cos(\omega t - kx + \phi). \tag{2.7}$$

The wavelength of any sinusoidal function is the distance it takes for the waveform to repeat a full cycle. The distance to the local pressure maxima is half the full cycle.

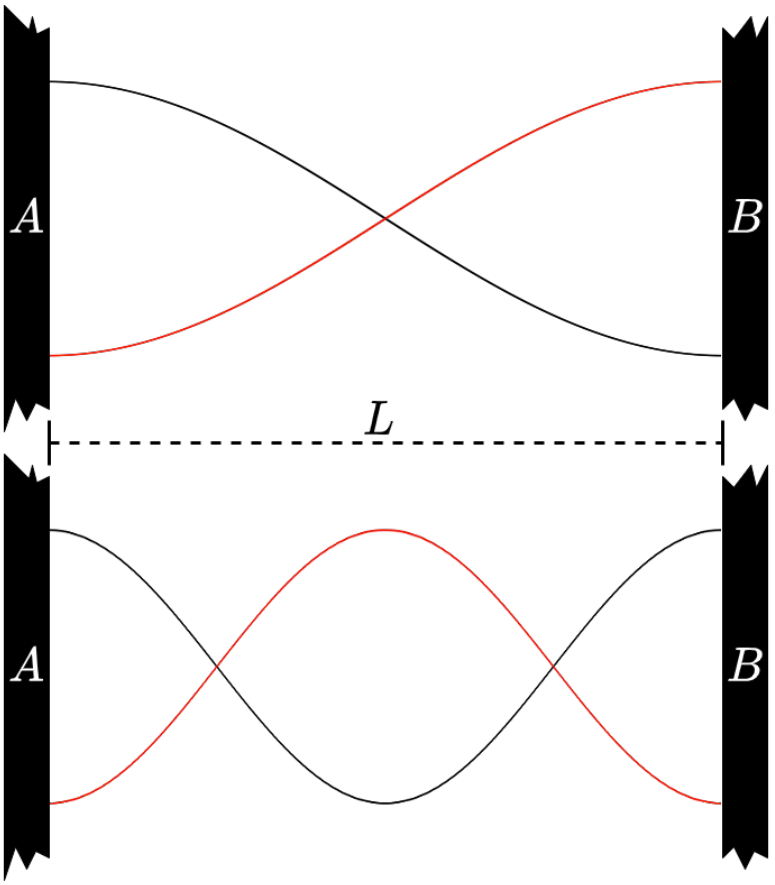


FIGURE 2.5: First and second room pressure modes shown as sinusoidal functions with the lines showing the displacement from the mean. — shows the propagation where $\cos(\omega t - kx + \phi) = 1$ at $x = 0$, and — shows the propagation where $\cos(\omega t - kx + \phi) = -1$ at $x = 0$

.

Figure 2.5 shows the same system as Figure 2.3, this time presenting the first two pressure modes as sinusoidal waves with wavelengths of $2L$ and L, respectively. Although

the standard graphical representation of these waves as shown in Figure 2.5 gives the impression that they are transverse waves, it is important to remember that they are longitudinal waves. In Figure 2.5, there are areas where the waves have maximum displacement from atmospheric pressure, i.e. anti-nodes, and areas where there is no displacement, i.e. nodes. In a theoretical system such as the one discussed in this section, a listener placed in a node would not hear sound at that frequency, no matter how powerful the source was. The real-world effect of these nodes and anti-nodes leads to the peaks and troughs of the room response shown in Figure 2.2, with the signal heavily attenuated in the nodes.

TABLE 2.1: Mode number to wavelength to frequency

Mode no.	Wavelength	Frequency
1	$\lambda_1 = 2L$	$f_1 = \frac{c_0}{2L}$
2	$\lambda_2 = L$	$f_2 = \frac{c_0}{L}$
3	$\lambda_3 = \frac{2L}{3}$	$f_3 = \frac{3c_0}{2L}$
n^{th}	$\lambda_n = \frac{2L}{n}$	$f_n = \frac{nc_0}{2L}$

Table 2.1 combines the work of this section, giving general expressions for calculating the wavelength and frequency of the modes of a one-dimensional space.

2.1.3.2 Three-dimensional modes

This section transfers the concept from one-dimensional space to a three-dimensional cuboid enclosure and does not apply directly to non-cuboid enclosures. Throughout this thesis a three-dimensional Cartesian coordinate system in the form $\mathbf{r} = (x, y, z)$ will be assumed unless otherwise stated, and the modal frequencies will be related to their modal numbers as $f_{(n_1,n_2,n_3)}$.

In three dimensions, the problem becomes more complex. The system is now enclosed, with six joined rigid boundaries leading to three possible types of mode:

- **Axial modes -** The path travelled between any two parallel surfaces, as discussed in the previous section. Only one of the mode numbers $f_{(n_1,n_2,n_3)}$ has an integer value, and the other two equal zero.
- **Tangential modes -** The path travelled between any two pairs of parallel surfaces. Two of the mode numbers in $f_{(n_1,n_2,n_3)}$ have integer values, and the third equals zero.
- **Oblique modes -** The path travelled between all three pairs of parallel surfaces. All three of the mode numbers in $f_{(n_1,n_2,n_3)}$ have integer values.

As seen in Table 2.1, the wavelength is proportional to twice the distance travelled. If $\mathbf{L} = (l_1, l_2, l_3)$ are the lengths of each boundary in the three-dimensional system, this distance can be calculated using de Gua's theorem (Conant and Beyer (1974)) to yield a wavelength of

$$\lambda_{(n_1,n_2,n_3)} = 2\sqrt{\left[\frac{l_1}{n_1}\right]^2 + \left[\frac{l_2}{n_2}\right]^2 + \left[\frac{l_3}{n_3}\right]^2}. \tag{2.8}$$

Therefore, the modal frequency of a three-dimensional system can be expressed as

$$f_{(n_1,n_2,n_3)} = \frac{c_0}{2}\sqrt{\left[\frac{n_1}{l_1}\right]^2 + \left[\frac{n_2}{l_2}\right]^2 + \left[\frac{n_3}{l_3}\right]^2}. \tag{2.9}$$

2.1.3.3 Calculation of modes in a three-dimensional enclosure

The work in Section 2.1.3 has been calculated from first principles using simple mathematics, and as such is perfectly valid. In this section, formulae are introduced in a textbook style, with less discussion of proof. Two methods of calculating the modes in a three-dimensional cuboid enclosure are shown. A comparison is then made between the second of these methods and real-world test data.

Taking. (2.9) and combining it with $\omega = 2\pi f$ and $\mathbf{n} = (n_1, n_2, n_3)$ leads to

$$\omega_n = c_0\sqrt{\left[\frac{n_1\pi}{l_1}\right]^2 + \left[\frac{n_2\pi}{l_2}\right]^2 + \left[\frac{n_3\pi}{l_3}\right]^2}, \tag{2.10}$$

taken from Nelson and Elliott (1992). Individual wave number values can be calculated as

$$k_n = \frac{\omega_n}{c_0}, \tag{2.11}$$

which are the eigenvalues of the three-dimensional homogeneous Helmholtz equation

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = 0. \tag{2.12}$$

For a cuboid room with rigid boundaries, (2.12) can be used to calculate the harmonic pressure fluctuation of a particular frequency in the enclosure, in which the complex pressure must satisfy the Helmholtz equation (Nelson and Elliott (1992)). In Equation 2.12,

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{2.13}$$

and

$$\psi_{\mathbf{n}}(\mathbf{r}) = \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\sum_{n_3=0}^{\infty} Q_n \cos(\frac{n_1\pi x}{l_1})\cos(\frac{n_2\pi y}{l_2})\cos(\frac{n_3\pi z}{l_3}). \tag{2.14}$$

Here, (x, y, z) are the coordinates of the listening position, and $\psi(\mathbf{r})$ is referred to as the eigenfunction, normally denoted as $\psi_{\mathbf{n}}$. Throughout this thesis, a compact notion will be used replacing

$$\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\sum_{n_3=0}^{\infty} \quad \text{with} \quad \sum_{\mathbf{n}=\mathbf{0}}^{\mathbf{N}} \quad \text{or} \quad \sum_{\mathbf{n}=\mathbf{0}}^{\propto},$$

where $\mathbf{n}$ denotes a trio of integers (n_1, n_2, n_3) that are the individual mode numbers for each dimension, $\mathbf{0}$ is a vector holding the lower bound of summation of such that $\mathbf{0} = (0, 0, 0)$, and $\mathbf{N}$ is a vector holding the upper bound of summation such that $\mathbf{N} = (N_1, N_2, N_3)$ or $\propto = (\infty, \infty, \infty)$. Equations 2.13 and 2.14 leads to Equation 2.12 being rewritten as

$$\nabla^2\psi_n(\mathbf{r}) + k^2\psi_n(\mathbf{r}) = 0. \tag{2.15}$$

Each eigenvalue is linked to an eigenfunction that gives the spatial pressure variance associated with a given mode. This method can now be used to calculate the modal frequencies and associated spatial pressure variance within a cuboid enclosure with rigid boundaries.

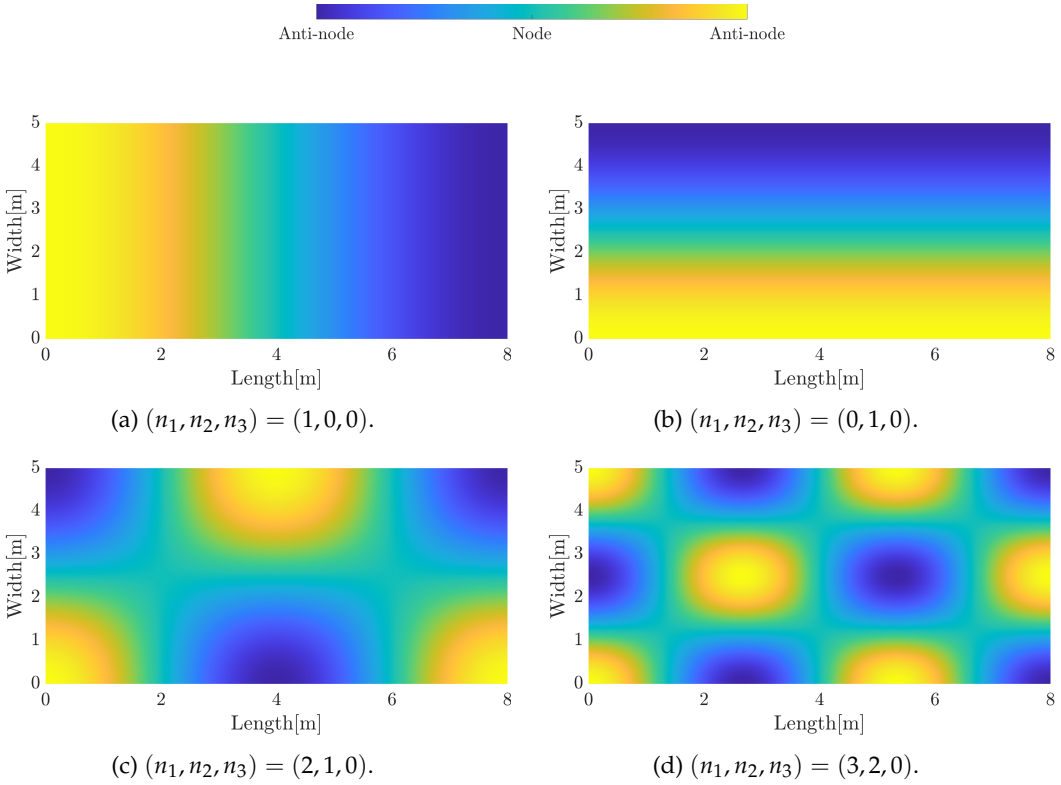
As an example, using (2.9), in the case that (l_1, l_2, l_3) = (8 m, 5 m, 2.7 m), the first ten modal frequencies are shown in Table 2.2.

TABLE 2.2: First ten modes of enclosure with dimensions (8 m, 5 m, 2.7 m)

(n_1, n_2, n_3)	Frequency	(n_1, n_2, n_3)	Frequency
(0,0,0)	0.00 Hz	(2,1,0)	54.90 Hz
(1,0,0)	21.44 Hz	(0,0,1)	63.52 Hz
(0,1,0)	34.30 Hz	(1,0,1)	64.31 Hz
(1,1,0)	40.45 Hz	(3,0,0)	67.03 Hz
(2,0,0)	42.88 Hz	(0,1,1)	68.60 Hz

It can be seen in Table 2.2, that with an enclosure of these dimensions, there are no oblique modes in the first ten resonance frequencies. The mode (0, 0, 0), known by several names, including the zeroth mode and the acoustic compliance mode, indicates the acoustic compliance of the volume of the enclosure. Using (2.15), the spatial pressure field within the enclosure can be calculated, with the values in Table 2.2 generating the eigenvalue. The eigenfunction can then be calculated for any values of (x, y, z).

Figure 2.6 shows how the acoustic pressure of four modes fill the enclosure. In all cases, there are anti-nodes (where the pressure fluctuates between maximum displacement



(a) $(n_1, n_2, n_3) = (1, 0, 0)$.

(b) $(n_1, n_2, n_3) = (0, 1, 0)$.

(c) $(n_1, n_2, n_3) = (2, 1, 0)$.

(d) $(n_1, n_2, n_3) = (3, 2, 0)$.

FIGURE 2.6: Pressure throughout the enclosure for four modes.

from the mean) and nodes (where there is no fluctuation) throughout the enclosure. Plots 2.6a and 2.6b show the first axial mode in the x and y planes respectively, while Plots 2.6c and 2.6d show two slightly more complex tangential modes. Although this can provide an idea of how acoustic pressure might fill the enclosure, this method of calculation lacks two important variables:

1. It considers there to be no source.
2. It models the walls as rigid boundaries with no absorption.

These issues can be addressed by using the Green's function. The Green's function is a solution to an inhomogeneous differential equation describing an impulse response (Sokolnikoff (1966)) and can be used to describe a wide range of physical phenomena (not only acoustics). A full analysis is given by Pierce (1981), and further information is also provided by Kuttruff (2017), Nelson and Elliott (1992), Williams (1999) and Morse and Ingard (1968). As the derivation of Green's functions is well-documented, there is no need to repeat the full details here. Briefly, the basic process begins by adding a monopole source to the enclosure at point $\mathbf{s} = (\tilde{x}, \tilde{y}, \tilde{z})$. This $\mathbf{s}$ is represented as a three-dimensional Dirac delta function, meaning the complex pressure at $\mathbf{r}$ resulting from $\mathbf{s}$, can be defined as

$$\delta(\mathbf{r}-\mathbf{s}) = \delta(x-\tilde{x})\delta(y-\tilde{y})\delta(z-\tilde{z}). \tag{2.16}$$

The Green's function can be indicated as $G(\mathbf{r}|\mathbf{s})$, which is the value at point $\mathbf{r}$ resulting from a monopole source at location $\mathbf{s}$, and this can be combined with Equation 2.16 to create the inhomogeneous Helmholtz equation

$$(\nabla^2 + k^2)G(\mathbf{r}|\mathbf{s}) = -\delta(\mathbf{r}-\mathbf{s}). \tag{2.17}$$

This must still satisfy the ridged boundary condition such that $\boldsymbol{\nabla}G(\mathbf{r}|\mathbf{s}).\mathbf{n} = 0$. It can then be stated that

$$G(\mathbf{r}|\mathbf{s}) = \sum_{\mathbf{n}=\mathbf{0}}^{\infty} \gamma_{\mathbf{n}}(\mathbf{s})\psi_{\mathbf{n}}(\mathbf{r}), \tag{2.18}$$

where $\gamma_{\mathbf{n}}(\mathbf{s})$ is a set of currently unknown complex coefficients whose values depend on the location of $\mathbf{s}$. From Equation 2.15, it can be shown that $\nabla^2\psi_{\mathbf{n}}(\mathbf{r}) = -k_{\mathbf{n}}^2\psi_{\mathbf{n}}(\mathbf{r})$, and this can be combined with Equation 2.17 and Equation 2.18 to yield

$$\sum_{\mathbf{n}=\mathbf{0}}^{\infty} \gamma_{\mathbf{n}}(k^2 - k_{\mathbf{n}}^2)\psi_{\mathbf{n}}(\mathbf{r}) = -\delta(\mathbf{r}-\mathbf{s}). \tag{2.19}$$

Nelson and Elliott (1992) shows how, from the starting point given in Equation 2.19 and with substitution for $\gamma_{\mathbf{n}}$ and integration over volume V, the Green's function can be expressed as

$$G(\mathbf{r}|\mathbf{s}) = \sum_{\mathbf{n}=\mathbf{0}}^{\infty} \frac{\psi_{\mathbf{n}}(\mathbf{r})\psi_{\mathbf{n}}(\mathbf{s})}{V(k_{\mathbf{n}}^2 - k^2)}, \tag{2.20}$$

where

$$\begin{aligned} \psi_{\mathbf{n}}(\mathbf{r}) &= \cos\frac{xn_1\pi}{l_1}\cos\frac{yn_2\pi}{l_2}\cos\frac{zn_3\pi}{l_3} \\ \psi_{\mathbf{n}}(\mathbf{s}) &= \cos\frac{\tilde{x}n_1\pi}{l_1}\cos\frac{\tilde{y}n_2\pi}{l_2}\cos\frac{\tilde{z}n_3\pi}{l_3}. \end{aligned} \tag{2.21}$$

This provides a method to calculate values at a location $\mathbf{r}$ resulting from a signal radiating from source location $\mathbf{s}$ but still models the walls as rigid boundaries with no absorption. Nelson and Elliott (1992) also shows that although absorption can be added, leading to an approximated model, this model assumes only a small amount of absorption. $\zeta_{\mathbf{n}}$ has been used to represent the damping coefficient to yield

$$G(\mathbf{r}|\mathbf{s}) = \sum_{\mathbf{n}=\mathbf{0}}^{\infty} \frac{\psi_{\mathbf{n}}(\mathbf{r})\psi_{\mathbf{n}}(\mathbf{s})}{V(k_{\mathbf{n}}^2 - k^2 + ikk_n c_0\zeta_{\mathbf{n}})}, \tag{2.22}$$

This can then be modified to give the radiated pressure at point $\mathbf{r}$ resulting from a monopole source at location $\mathbf{s}$, such that $\hat{\mathrm{P}}(\mathbf{r}|\mathbf{s}) = i\rho_0 c_0 k Q \mathrm{G}(\mathbf{r}|\mathbf{s})$ (Nelson and Elliott (1992))

$$\hat{\mathrm{P}}(\mathbf{r}|\mathbf{s}) = \sum_{n=0}^{\infty} \frac{\omega\rho_0 c_0^2 Q\ \psi_n(\mathbf{r})\psi_n(\mathbf{s})}{[2\zeta_n\omega_n\omega + i(\omega^2 - \omega_{\mathbf{n}}^2)]\mathrm{V}} \tag{2.23}$$

This is how the Green's function will be defined throughout this thesis, providing the radiated pressure at any point $\mathbf{r}$ resulting from any source $\mathbf{s}$, where Q is the volume velocity (within the scope of this thesis, $Q = 1$ unless otherwise stated), and $\zeta_{\mathbf{n}}$ is the damping coefficient.

The damping coefficient should be representative of the real-world conditions; therefore, it should be calculated from the RT_{60}. Based on the work by Nelson and Elliott (1992), it can be seen that the impulse response of a single mode has an envelope of the form $e^{-\zeta_n\omega_n t}$. This can be expressed in terms of the RT_{60} as

$$e^{-\zeta_n\omega_n t} = \begin{cases} 1 & t = 0 \\ 10^{-3} & t = \mathrm{RT}_{60} \end{cases} \tag{2.24}$$

therefore

$$\zeta_n = \frac{3\ln(10)}{\mathrm{RT}_{60}\omega_n}. \tag{2.25}$$

In Section 2.1.1, it was stated that the target RT_{60} should be between 0.2 and 0.5 seconds from 250 Hz to 4 kHz. Additionally, this target may rise by 1.25% at low frequencies. This would give a RT_{60} range between 0.25 and 0.625 seconds, which, using Equation 2.25, yields ζ_n values of $\frac{27.631}{\omega_n}$ and $\frac{11.052}{\omega_n}$, respectively. This RT_{60} method for defining the damping will be used in the analytical models in Chapters 3 and Chapters 4.

The work in Chapter 5 relies on the use of Green's function to replace real-world measurements. Textbooks such as Kinsler et al. (1982), Kuttruff (2017), Morse and Ingard (1968), Nelson and Elliott (1992), Pierce (1981) and Williams (1999) give excellent full mathematical proofs of Green's functions as solutions to inhomogeneous acoustic differential equations, and state that it provides a good approximation of real-world measurements. However, the data comparing Green's functions to the real world is sparse. Two papers, Luan and Jacobsen (2008) and Cox et al. (2004), demonstrate that Green's functions can provide good estimations of real-world measurements, but the authors do not discuss any methods for damping estimation. The most detail is provided by Luan and Jacobsen (2008), who demonstrated using a precise measurement method where the loudspeaker was mounted inside an inverted horn, with two matched microphones mounted at the aperture. This, in a lightly damped room with dimensions

(3.29, 4.38, 3.29 m), showed that the measurement taken at a third microphone matched the theoretical Green's function below 320 Hz. The main issue with this approach is that such a precise measurement method would not be available on the home market. More generally, it may be that the number of factors that would need to be taken into account for a real-world, at-home setting is limiting the work in this area. Examples of such factors are the accurate measurement of room dimensions, the locations within the space and the damping coefficients of each surface, whether they are uniform across the surface and how they change with frequency.

As shown in Figure 2.7, the RT_{60} method provides an approximate real-world measurement due to the low modal density in the frequency band of interest when calculating RT_{60} (Jambrošić et al. (2008)). In Chapter 5, various methods will be compared to attempt to improve the damping calculations, such that Equation 2.23 can provide a more accurate estimation of real-world measurements. The methods used for estimating the damping are based on existing work. However, the integration of damping calculations within Green's function should be considered a novel development. Two methods have been introduced with the aim of estimating the damping of the individual low-frequency modes. The first method estimates the mode's Q factor from frequency response data and is well documented by Noxon (1986) and many others, including Kinsler et al. (1982), Kuttruff (2017), Morse and Ingard (1968), Everest (1989) and Pierce (1981). The second method estimates the individual low-frequency mode's decay times and is referred to as the MT_{60} method. It is derived from Magalotti and Cardinali (2018), Magalotti and Cardinali (2019), Magalotti (2019) and Magalotti and Ponteggia (2019). Both these methods will be fully discussed in Chapter 5, and are only mentioned here because data from a variant of the MT_{60} method is shown in Figure 2.7.

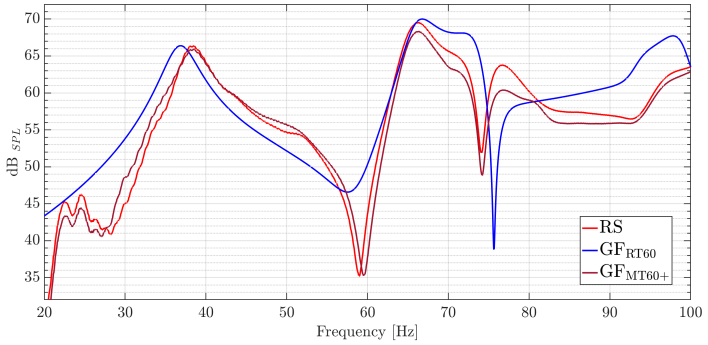


FIGURE 2.7: Comparison of frequency response at a single receiver from a single source with real-world measurements (—RS) using two analytical models: the RT_{60} method (—GF_{RT60}) and the MT_{60+} method (—GF_{MT60+}).

Figure 2.7 shows an example taken from the work in Chapter 5. In the present section, it is used to demonstrate how the Green's function can estimate a real-world measurement. As shown in Figure 2.7, the RT_{60} method provides an approximate estimation of the real-world measurement. Some of the main trends and shapes can be seen, but

it cannot be called accurate. The other method, shown here as MT_{60+}, is essentially a Green's function with mathematical enhancement used to better reflect the environment. The modelling it achieves is much closer to the real-world measurement and demonstrates that it is possible for Green's function to provide an accurate estimation of real-world measurements.

2.1.4 Section summary

This section has provided the necessary background knowledge on acoustics to understand the remainder of this work. It has shown that within an enclosed listening environment, two main issues affect audio.

1. **Dimensions -** These define the resonance frequencies.
2. **Damping -** This defines the length of time that the resonance frequencies reverberate for.

The Green's function has been introduced as an analytical model and will be used throughout this thesis. Furthermore, it has been shown that with some mathematical enhancement (the details of which are explained in Chapter 5), Green's function can provide a result that is close to real-world measurements.

2.2 Control of room acoustics

This section reviews previous approaches for controlling room acoustics, focusing on two main control methods: digital room correction and acoustic cancellation. Although not the focus of this work, it is important to mention that there are also a number of highly effective passive control methods. However, these methods require physical changes to the listening environment through the addition of materials or structures that absorb or diffuse acoustic energy. In some situations, décor can act as passive control and soft furnishings will absorb sound waves. There is a large body of research on this, including Everest (1989) and Kuttruff (2017).

In this thesis, the terms digital room correction and acoustic cancellation are defined as:

- **Digital room correction -** A signal processing method that filters audio to compensate for the enclosure's frequency response before broadcasting from the loudspeaker.
- **Acoustic cancellation -** The controlled cancellation of acoustic pressure within an enclosure.

There is some crossover between the two methods. For example, a digital room correction method may result in or rely on some acoustic cancellation, or an acoustic cancellation approach may include minimal signal processing to adjust the signal level and delay, before broadcasting the loudspeaker's audio into the enclosure. However, in this section the methods have been separated based on their primary approaches. Each of the two overarching approaches accounts for many previously proposed systems, with specific targets and predefined metrics to assess how well they achieve those targets. In simple terms, the goal of most systems is to create a relatively flat frequency response throughout a listening area and ensure the modal resonances are evenly spread out so there are no overly dominant modes at any given frequencies.

An excellent review of many of these methods and options is provided by Cecchi et al. (2017). The work in this section aims not to repeat Cecchi et al. (2017), although there is a high level of crossover.

2.2.1 Digital room correction

Digital room correction approaches often include varying the number of loudspeakers. The location of each loudspeaker is not prescribed, and the aim of all approaches is to control one or more listening zones. However, once the filters are designed, the system variables cannot be changed without recalculating the filters. Digital room correction is known by several names, including de-reverberation and deconvolution, and it has been covered in many papers, including Mourjopoulos (1994), Karjalainen et al. (2001), Elliott and Nelson (1989) and Norcross et al. (2006). As discussed in Section 2.1, an acoustic measurement considers the combination of the audio and the room effect at a specific location. The room effect can be extracted from a measurement, and a filter with the inverse response can be created and applied to the audio before broadcasting from the loudspeaker. This simple theory assumes that the filter and the room will cancel each other out at the point of measurement in the enclosure, leaving just the audio. The number of sources can be increased because the audio at the listening position is the linear superposition of all the individual sources. Additionally, each source will have its unique room response and inverse filter.

There are a few obvious issues with this method. First, it has been discussed in Section 2.1 that the measured frequency response of any environment depends on the location of the source and the receiver. Therefore, this inverse filter is only valid for the exact locations of source and receiver at the time of measurement (and the exact conditions in the environment at that time, including the temperature and number of people present). Even if the source locations are fixed, the listener location may not be, and there may also be more than one listening position. This filter technique may yield unpredictable and even unpleasant results in such locations. Secondly, if the measurement is taken within a node, the system may attempt to add large amounts of gain

to the nodal frequency, achieving little (or no) improvement at the listening position, while making this frequency very loud throughout the rest of the environment. These issues are commonly addressed by acquiring several measurements across a listening area, then comparing and averaging these measurements to create the filter. Accordingly, there are two main categories of measurement system:

1. **Single-point measurement** - Single-point measurement options estimate the filters based on the measurement in a single location. This is outlined by Neely and Allen (1979) and is shown to be effective only in a limited zone around the measurement point.

2. **Multiple-point measurement** - In multiple-point measurement, multiple measurements are taken either across defined listening areas or across the whole space. Typically, some form of spatial averaging is then used to create filters. These systems may not offer the level of control of a single-point measurement system at any single point, but they aim to provide a more balanced improvement. They differ in terms of how they group the measurements, what they target, how they rate effectiveness and how they design and apply the filters.

These systems can both have either

1. **A single source** - One loudspeaker offering a single degree of freedom.

2. **Multiple sources** - Multiple loudspeakers offering multiple degrees of freedom. These systems are normally more effective than those with a single source as, with each source having its own filter, the sources can work as a group to obtain the best overall improvement (Elliott and Nelson (1989)).

The main focus of this thesis is multiple-point measurement systems with multiple sources. However, before reviewing the measurement grouping methods for multiple-point measurement systems, there is one multiple-point measurement system that does not require averaging. The multiple-input/multiple-output inverse theorem was put forward by Miyoshi and Kaneda (1986) and Miyoshi and Kaneda (1988). This approach requires that the number of measurement locations be less than or equal to the number of sources. This yields either an underdetermined or fully determined matrix of acoustic measurements with an inverse mathematical solution that will allow for the creation of perfect inverse filters to correct for the issues (or impose targets) at the measurement points. However, due to the required relationship between the number of sources and receivers, it was decided that this system was impractical for the work in this thesis.

2.2.1.1 Measurement grouping

In almost all practical multiple-point measurement systems, numerous measurements are taken over a single or multiple listening areas, which the system attempts to control. The fact that there are more measurement points than sources leads to an overdetermined matrix of acoustic measurements that may have no inverse mathematical solution. This means that any control will always have to be averaged across the set of measurement points to achieve as close a representation of its target as possible.

One of the earliest approaches for averaging across a set of measurement points, introduced by Elliott and Nelson (1989), is based on the design of a set of digital equalisation filters to average across the points. The method outlined by Elliott and Nelson (1989) shows how to build equalisation filters using the least-mean-squares approach, meaning the filter coefficients are calculated to minimise the sum of the squared errors between the transfer function and target, across all measurement points. Elliott and Nelson (1989) also showed that this method works well for both single-source and multiple-source systems. However, as discussed in both Bharitkar et al. (2002) and Bharitkar et al. (2003), using the least-mean-squares approach for averaging over a set of measurement points often includes the major limitation that averaging treats all points as equal. Thus, it cannot exploit or focus on any similarities in room responses, nor can it favour equalisation at certain positions to solve some of the issues with the average and weighted average method. To address this issue, Mourjopoulos (1994) introduced the clustering method, which takes a large number of transfer functions and groups them by similar characteristics. These groups are then used to create correction filters, which initiates a recursive phase where cluster sizes (and thus filters) are adjusted to provide the best response.

Bharitkar and Kyriakakis (2001) introduces the concept of fuzzy c-means clustering. In the original clustering method, a transfer function can only belong to one cluster, whereas in fuzzy c-means clustering, any transfer function can belong to multiple clusters. Each transfer function in a cluster can have a unique weight, which denotes the importance of that function to the cluster. For example, in the normal clustering (hard c-means clustering) method, transfer functions with a peak at 50 and 100 Hz are clustered together, ignoring transfer functions that have a peak only at either 50 or 100 Hz. In fuzzy c-means clustering, these transfer functions would be included in the cluster but given lower scores. The latter clustering methods have their own drawbacks; for example, when addressing specific issues (at specific locations or frequencies), the clustering may not improve the entire listening area on average. This can be counteracted by using the original method from Elliott and Nelson (1989) as a limit, which the system will revert to should there be no average improvement when clustering.

Returning to the work as outlined by Elliott and Nelson (1989), Kirkeby and Nelson (1993) uses the least-mean-squares approach for acoustic pressure matching to reproduce a plane-wave sound field and asks whether it is possible to perfectly reproduce low-frequency sound in the listening space. It is elegantly demonstrated that the quality of the reproduced sound can never be completely perfect, and further postulated that the quality can be determined by the size of the receiver area and the angles between the sources of the recorded sound from the perspective of the middle of the receiver array. In later chapters, this thesis will build on the work by Kirkeby and Nelson (1993) and Elliott and Nelson (1989) as it is anticipated that the least-mean-squares approach will provide the best overall method of control for the sound field under investigation, with the possibility of clustering added by subsequent work, once the underlying theory has been proven.

Although there are issues with the average or weighted-average method systems, it is the most widely found method in the consumer market. In this section, a number of the leading systems within the market will be discussed. All of the systems work in a similar way: measurements are taken, after which the system creates a room correction set of filters based on an algorithm. It is difficult to speculate on the exact mathematical process used in each consumer system, as these are normally part of the company's intellectual property and unique selling proposition. This means that, unless patented, the details are often not in the public domain. The standard outcome is either a finite impulse response (FIR) or infinite impulse response (IIR) filter. Many room correction systems try to use minimum phase filters (Mourjopoulos (1994)), and these systems also tend to focus on a section of the audio bandwidth, which can allow for higher-precision filters to be designed so as not to waste processing power on controlling unnecessary frequencies.

TABLE 2.3: Comparison between commercial digital equalisation systems

Manufacturer	System	Maximum Channels	Frequency Range	Maximum Correction	Filter Type
Audyssey	MultiEQ	5.1	Full band	±12 dB	FIR
B & O	ABC	2	20 -500 Hz	± 10 dB	***Unknown***
Copland	DRC 205	2	Full band	***Unknown***	***Unknown***
Deqx	PDC 2.6	2	Full band	***Unknown***	Parametric + FIR
Direc	Live	5.1	Full band	***Unknown***	***Unknown***
Meridian	861 Processor	7.1	20 - 250 Hz	-20 to 0 dB	IIR

Table 2.3 presents a simple comparison between the systems. The Audyssey and Deqx use FIR filters. It is widely accepted in digital signal processing that FIR filters are often easier to design than IIR filters, as the coefficients can be calculated directly from the impulse response in the time domain (Smith (2003)). However, FIR filters are normally more processor-intensive as they require more coefficients for a given steepness

of the filter than IIR filters do (Johansson (2007)). Normally, to control a specific frequency, an FIR filter also needs to be built from a time-domain impulse response, equivalent in length to that frequency's wavelength divided by the speed of sound (Smith (2003)). Furthermore, the more processor-intensive the algorithm is, the higher the cost of the digital signal processor (DSP) chip will be. In contrast, IIR filters tend to be less processor-intensive, requiring fewer coefficients but being more mathematically complex in design (Johansson (2007)). For example, adding a notch filter at 20 Hz on a DSP running at a sample rate of 8 kHz (this being the lowest sample rate at which most low-cost DSP chips in the market can currently run natively), would necessitate an impulse response of 400 samples/coefficients. This requires the DSP chip to run 400 instructions per cycle of audio processing per channel to apply the needed filters. The 400 instructions per cycle could provide 80 second-order IIR filters. A good approximation of a 20-Hz notch filter with an IIR filter requires one filter with five coefficients and five instructions per cycle. In the Meridian whitepaper 861v8 DS v2.7 (2012), it is stated that the original 861 Processor had 300 instructions per cycle per channel for all the required processing, which limits the length of the FIR filters. Various methods can be used to reduce the FIR filter length, and a single FIR filter can replace multiple IIR filters to achieve a balance between performance and cost. The manufacturer has to make a decision on this point when choosing which filter methods to use and whether, and/or how, they are going to combine the two.

As shown in Table 2.3, the B & O and Meridian systems limit the bandwidth over which control is applied, allowing the DSP to focus all of its power on a narrow band. Although this can help with processing, a detrimental effect on the absolute phase and latency between the processed and unprocessed sections of the spectrum must be taken into consideration. Due to these systems' commercial nature, it can be difficult to find details on how the data is processed. However, some details are available for the Audyssey and B & O approaches.

- The Audyssey system, as described by Bharitkar and Kyriakakis (2011) and Bharitkar (2003), is a multiple-point measurement approach that can work with both multiple-source and single-source systems. The system requires several microphone measurements to be made (from 8 to 32) in non-uniform placements across the listening area. The system creates a set of weighted averages based on a psycho-acoustic model, which are grouped to determine the best outcome using three methods: hard c-means clustering, fuzzy c-means clustering and adaptive learning with linear predictive coding (LPC) and pole-zero.

- The B & O system is a two-point measurement approach with a single source, but it requires that the two measurements be made in two different layouts. First, the loudspeaker and microphone are placed in the reference positions, which are to be defined by the user as the ideal positions within the environment. Then, both

are placed in their normal positions. Next, a filter that imposes the response taken at the reference positions onto the normal positions is created (Pedersen (2012)).

Two other methods of measurement grouping include common acoustic poles compensation (Haneda et al. (1994)) and modal equalisation (Karjalainen et al. (2001)). First proposed by Haneda et al. (1994) and Haneda et al. (1997), common acoustic poles compensation aims to estimate and control the poles (modes) of an environment, independent of the source and receiver position. The acoustic poles are estimated from transfer functions or the multiple-point measurements, using an LPC model of the room response. The equalisation is then achieved with an FIR (all-pole) inverse filter, which has the inverse characteristics of the common acoustic pole function. Karjalainen et al. (2001) defines modal equalisation as a process that can modify the rate of modal decay. The natural decay of a mode is dependent on the absorption of the walls between which the standing wave is created. To shorten the decay time, the Q factor of the resonance needs to be decreased, which is achieved by moving the filter pole toward the origin in the z-plane. Modal decay time modification can be implemented on the sound entering a room either by filtering the source or introducing additional control sources to interact with the primary sound.

2.2.2 Acoustic cancellation

Acoustic cancellation, also known as active absorption, has been an area of ongoing interest for many researchers, including Celestinos and Nielsen (2004), Celestinos and Nielsen (2005), Celestinos and Nielsen (2006), Celestinos and Nielsen (2008a), Celestinos and Nielsen (2008b), Santillán (2001), Santillán et al. (2007), Nelson and Elliott (1992), Vanderkooy (2007) Vanderkooy (2011) and Vanderkooy and Rousseau (2013). Acoustic cancellation approaches involve one or more primary sources that broadcast an audio signal into the environment and then use one or more secondary sources to broadcast a secondary audio signal into the environment. The interaction between the two signal types results in a net reduction of the acoustic pressure in the environment. Two acoustic cancellation approaches are discussed in this section:

1. Plane wave cancellation
2. Spherical wave cancellation

2.2.2.1 Plane wave cancellation

The papers Santillán (2001) and Santillán et al. (2007) were some of the first to introduce, outline and test the plane wave cancellation method as a solution to the acoustic

issues related to listening rooms, as outlined in Section 2.1. This work was continued in Celestinos and Nielsen (2004), Celestinos and Nielsen (2005), Celestinos and Nielsen (2006) and Celestinos and Nielsen (2008a), culminating in the controlled acoustic bass system (CABS) outlined by Celestinos and Nielsen (2008b). The central premise of all plane wave cancellation methods is that in a cuboid room where the longest dimension runs from the front to the rear of the room, a plane wave that will propagate down the room's length can be created by an array of sources on the front wall. This wave can then be cancelled at the room's rear with an identical array of audio sources on the rear wall. Both arrays broadcast the same signal. However, before broadcasting the audio from rear sources into the enclosure, minimal signal processing is used to adjust the signal level and take into account any loss in energy from the plane wave as it moves down the room with the phase of the signal inverted. Next, the signal is delayed by an amount proportional to the room's length. Unlike equalisation techniques that emphasize a single or small number of listening zones, the plane wave cancellation method covers the entire room (Celestinos and Nielsen (2008b)).

In Celestinos and Nielsen (2008b), CABS was both modelled in a simulation and assessed with real-world measurements in two listening rooms. One room conformed to the International Telecommunication Union (ITU) standard (ITU-R:BS.775-3 (2012)), the other to the International Electrotechnical Commission (IEC) standard (IEC:60268-13 (1998)). The main tests of interest are the setups the paper denotes as CABS 0.2.0 and CABS 0.2.2, which involve two and four sources, respectively. The CABS 0.2.0 setup has no cancellation, with the two sources set up at the front of the room. CABS 0.2.2 is a plane wave cancellation setup with four sources, the front two sources being the same as in CABS 0.2.0, with two additional sources at the rear of the room. Defining the rooms length, width and height as (l_1, l_2, l_3), the dimensions are given in Table 2.4, and defining the four source locations as s_1 to s_4 the locations of the sources in each study are provided in Table 2.5.

TABLE 2.4: Room dimensions for CABS simulation data and real-world measurements taken from Celestinos and Nielsen (2008b).

Room	Dimensions (l_1, l_2, l_3)
simulation	(7.80 m, 4.20 m, 2.76 m)
IEC	(7.80 m, 4.12 m, 2.78 m)
ITU	(8.13 m, 7.39 m, 2.88 m)

The data measurements were taken using a five-by-five receiver grid placed in the centre of the room, and for the analysis, the metric defined in Celestinos and Nielsen (2008a) as the mean sound field deviation was used. This metric comprised two values. First, the spatial deviation (SD) over the frequency band of interest was defined as

TABLE 2.5: Source locations for CABS simulation data and real-world measurements taken from Celestinos and Nielsen (2008b).

Source No.	Source Location (l_1, l_2, l_3)		
	simulation	IEC	ITU
s_1	$(0.06\text{ m}, \frac{l_2}{4}, \frac{l_3}{2})$	$(0.06\text{ m}, \frac{l_2}{4}, 1.50\text{ m})$	$(0.09\text{ m}, \frac{l_2}{4}, 1.44\text{ m})$
s_2	$(0.06\text{ m}, \frac{3l_2}{4}, \frac{l_3}{2})$	$(0.06\text{ m}, \frac{3l_2}{4}, 1.50\text{ m})$	$(0.09\text{ m}, \frac{3l_2}{4}, 1.44\text{ m})$
s_3	$(l_1 - 0.06\text{ m}, \frac{l_2}{4}, \frac{l_3}{2})$	$(l_1 - 0.06\text{ m}, \frac{l_2}{4}, 1.50\text{ m})$	$(l_1 - 0.09\text{ m}, \frac{l_2}{4}, 1.44\text{ m})$
s_4	$(l_1 - 0.06\text{ m}, \frac{3l_2}{4}, \frac{l_3}{2})$	$(l_1 - 0.06\text{ m}, \frac{3l_2}{4}, 1.50\text{ m})$	$(l_1 - 0.09\text{ m}, \frac{3l_2}{4}, 1.44\text{ m})$

$$SD = \frac{1}{N}\sum_{n=1}^{N}\sqrt{\frac{1}{M-1}\sum_{m=1}^{M}(L_{p_m}(f_n) - \overline{L}_{p_m}(f_n))^2}, \tag{2.26}$$

where L_{p_m} is the sound pressure level in dB at receiver points $\mathbf{r}_m$. Second, the magnitude deviation (MD) over the frequency band of interest was defined as

$$MD = \frac{1}{M}\sum_{m=1}^{M}\sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(L_{p_m}(f_n) - \overline{L}_{p_m}(f_n))^2}. \tag{2.27}$$

Based on Celestinos and Nielsen (2008b), it is clear that the plane wave cancellation method of sound control has some promise. This is backed up by both the simulation and experimental data, a sample of which is given here in Table 2.6 and Table 2.7.

TABLE 2.6: CABS simulation data taken from Celestinos and Nielsen (2008b) .

Setup	MSFD	
	SD	MD
CABS 0.2.0	±4.9	±6.6
CABS 0.2.2	±0.7	±1.9

TABLE 2.7: CABS experimental data taken from Celestinos and Nielsen (2008b). In the IEC room, the frequency band of interest was 20–100 Hz, and in the ITU room the frequency band of interest was 20–90 Hz

Setup		MSFD	
		SD	MD
IEC	CABS 0.2.0	±4.6	±6.4
	CABS 0.2.2	±1.6	±2.1
ITU	CABS 0.2.0	±4.2	±5.3
	CABS 0.2.2	±1.3	±2.1

All the data in Table 2.6 and 2.7 reveal a marked improvement in both the SD and MD when using the CABS 0.2.2 setup (plane wave cancellation method) over the CABS

0.2.0 setup (no-cancellation method). In both the SD and MD, the simulation exhibited a greater improvement than the experimental data; however, this was to be expected as the simulation will not be susceptible to unknown external variables. The delta in the SD values from the simulation was 4.2 dB. The delta dropped to 3.0 dB for the IEC room and 2.9 dB for the ITU room. The same trend can be observed in the MD delta as, again, the best real-world improvement was in the IEC room with a delta of 4.3 dB compared with the ITU room at 3.2 dB. Clearly, the dimensions of the space affect the results and, as stated by Celestinos and Nielsen (2008b), the ITU room had more damping, which may have caused an issue resulting from the creation of a plane wave. Additionally, the wider distance between the sources in the ITU room limited the upper frequency at which a plane wave could be created.

It is evident from both the simulation and experimental results in this work, that the CABS 0.2.2 setup (plane wave cancellation method) is an effective method for achieving optimum low-frequency sound pressure level distribution inside a cuboid room. It achieved good results over most of the room, contrasting with the single or small number of listening zones to which digital room correction is normally limited.

In addition to the work done by Celestinos and Nielsen (2008b), Fazenda et al. (2012) investigated the optimal listening setup using qualitative means. Here, eight different systems were set up in a room at the same time, thus allowing for a direct comparison. Twenty listeners were asked to choose their favourite arrangement using the paired comparison technique. Overall, it was found that the CABS and single source-to-sink (SSS) setups had the highest perceived quality (Fazenda et al. (2012)). The SSS setup can be thought of as a simplified version of CABS with one source at the front of the room and one at the rear, both at the halfway point of the width. This study had comparatively few participants (Martínez-Mesa et al. (2014)), but the experiment still has merit. Although it should not be implicitly assumed that the CABS and SSS setups are the best for home listening, there can be confidence that they offer a high level of subjective improvement to the audio experience. The CABS approach requires exact measurement and placement of the loudspeaker, as moving the loudspeakers away from the defined layout will impact the listening experience (Celestinos and Nielsen (2008b)). Additionally, if the room furniture arrangement changes, the cancellation effect of the two loudspeakers may be reduced due to additional diffraction. Overall, the CABS method is good at controlling low frequencies in listening rooms, but is quite inflexible for those inhabiting the space.

2.2.2.2 Spherical wave cancellation

The work of Celestinos and Nielsen (2008b) shows that plane wave cancellation offers a possible solution to the acoustic issues found with listening rooms as outlined

in Section 2.1. However, it does require fixed loudspeaker positions, and further testing would be needed to understand the extent to which this method would work in non-cuboid environments. A more useful system would give the same level of control without creating a plane wave. This section reviews work where spherical waves have been used, as outlined by Nelson and Elliott (1992) and extended in three papers: Vanderkooy (2007), Vanderkooy (2011) and Vanderkooy and Rousseau (2013). The main advantage of the spherical wave (point source) system is that it poses no constraints on the layout of loudspeakers within the environment.

The idea begins by using a point source absorber within a plane wave field to absorb power from the plane as it passes that point in the environment, leading to a net power reduction in the system as a whole. Vanderkooy postulates the idea that spherical sources can be used, showing that the mathematics leads to the same conclusion as Nelson and Elliot's theoretical work on plane wave sources.

The results from Vanderkooy (2011) and Vanderkooy and Rousseau (2013) are mixed. The finite element model shows promise, with the first two modes being reduced in level by approximately 20 dB and 10 dB, respectively. However, real-world testing does not support the modelled data and even shows an increase in the power at low frequencies rather than a decrease.

One issue is that the active absorber output needs to be a perfect inverse of the superposition of all sound fields to be cancelled within the environment at the point of cancellation. This can be achieved by placing a microphone in the centre of the pressure absorber's loudspeaker cone and using its output to create the signal to drive the absorber. This signal requires processing to remove the absorber's own output and create the inverse of the source signal at this single point. There will always be a delay in such a system, meaning the output will not be a perfect inversion. Thus, although it may cancel some power, it is more or at least as likely to add power to the environment. This area of study lies outside the remit of the present thesis, and further work could be done to determine whether the latency issue can be improved. An inverse filter could be designed beforehand from measurements taken at the appropriate location, and the loudspeaker feeds could be adjusted so that the time difference between the output of both source and absorber was as required. This would only be successful if the environmental conditions remained the same as during the measurements, including the number of people in the enclosure and their locations within it.

2.2.3 Section summary

This section has discussed methods for controlling the acoustic field inside enclosures. The extent to which the plane wave cancellation method can improve the audio experience both objectively and subjectively has been outlined, while its limitations have

also been introduced. The aim of Chapters 3 and 4 is to use ideas from digital room correction to put forward an approach for controlling the acoustic field inside enclosures that is objectively as good as the plane wave cancellation method, but without its limitations. To achieve this, Chapter 3 further investigates the plane wave cancellation method, discussing the theory that underpins the method and introducing studies that aim to understand more about the limitations of the method. Chapter 4 then uses a least-squares approach to design digital filters for room correction, and carries out an in-depth study comparing this with the plane wave cancellation method.

2.3 Metrics

Throughout this thesis, it is necessary to analyse the effectiveness of the acoustic control offered by each method and to describe various aspects of the sound field. Several important metrics are introduced here, which will be used and refined throughout Chapter 3. The aim is to create a set of critical metrics to achieve the most effective method of performing each subsequent analysis. Two of the metrics used in this work were introduced by Welti and Devantier (2006) and Welti (2012). Diverging from methods used by other researchers to optimise low-frequency response in living spaces, Welti investigated the possibility that simply changing the position where one sits in a room can dramatically affect the quality of the low-frequency response. Several factors profoundly impact the resultant low-frequency response, including the positioning of subwoofers and users in small spaces and the room's dimensions (Welti (2012)). While it is not feasible to state that one particular subwoofer configuration gives the best frequency response results, Welti posits that there is a way to estimate optimal configurations (Welti (2012)). As part of this work, Welti defines several metrics, of which this thesis will use two, the mean spatial variance (MSV) and spatial average (SA), which are detailed below. These aim to provide a simple objective method for comparing the effectiveness of the acoustic control. Welti also suggests that, although it has not been proven, certain metrics, such as mean spatial variance, correlate with subjective listener preference. If this were the case, the aim of minimising seat-to-seat variation would be a reasonable goal for any control system.

Spatial average. The spatial average (SA) is the average of the amplitude in dB between the measurements taken at M receiver locations at a given frequency, defined as

$$SA(f) = \frac{1}{M} \sum_{m=1}^{M} 20 \log_{10}[p_m(f)], \tag{2.28}$$

where $p_m(f)$ are the pressures at receiver points $\mathbf{r}_m$. This thesis uses a plot of SA(f) against frequency to replace the need for multiple-frequency responses from multiple-measurement points in the same environment. The SA(f) plot offers a rough visual guide of how close any method under test is to the target.

Mean spatial variance. The mean spatial variance (MSV) measures the receiver-to-receiver variance of the amplitude response. First, the spatial variance (SV) is defined as

$$SV(f) = \frac{1}{M-1} \sum_{m=1}^{M} (20 \log_{10}[p_m(f)] - SA(f))^2, \tag{2.29}$$

which is the variance of the amplitude in dB between the measurements taken at the receiver locations. The MSV over the frequency band of interest is then defined as

$$MSV = \frac{1}{N} \sum_{n=1}^{N} SV(f_n). \tag{2.30}$$

This gives the variance of sound level in dB as a single value, which is a function of the seating location (Welti (2009)), and if, as Welti also suggests, the MSV correlates with subjective listener preference, the lower the MSV, the more effective the system under test. From Welti (2012) it can also be postulated that the MSV offers three ranges:

1. **Low-level MSV** $<$ **15** : The variation among receivers demonstrates good system performance.
2. **Mid-level 15** $\leq$ **MSV** $\leq$ **35** : The variation among receivers falls within an acceptable range, suggesting satisfactory system performance.
3. **High-level MSV** $>$**35** : The variation among receivers surpasses the acceptable threshold, rendering the system unable to produce satisfactory results.

Throughout the thesis, the effectiveness of the system under test can be assessed firstly by the ranges in which the MSV falls and then by its MSV value.

Error function. When aiming to minimise the sum of squared errors between the transfer function and the target across all measurement points, an error function (e) is defined to express how close the result is to the target (Elliott (2000)). When comparing data sets in this thesis, up to three types of pressure signals can be measured at points denoted as $\mathbf{r}_m$. These can be defined as:

- $\mathbf{p}(f)$ - frequency vector of pressure at receiver points when no acoustic control or filter is used.
- $\mathbf{c}(f)$ - frequency vector of pressure at receiver points when any form of acoustic control or filter is used.
- $\mathbf{d}(f)$ - frequency vector of target pressure at receiver points.

therefore, two error functions can be defined:

$$e_p(f) = 10\log_{10}\frac{||\mathbf{d}(f)-\mathbf{p}(f)||^2}{||\mathbf{d}(f)||^2} \qquad \& \qquad e_c(f) = 10\log_{10}\frac{||\mathbf{d}(f)-\mathbf{c}_(f)||^2}{||\mathbf{d}(f)||^2}. \quad (2.31)$$

In the above equations, $e_p(f)$ is the error function when no form of acoustic control or filter is used, and $e_c(f)$ is the error function when any form of acoustic control or filter is in use. This thesis uses a plot of this error function against frequency to compare the effectiveness of the acoustic control methods for a given room with a set target.

The above analysis method does not allow for cross-comparisons between different rooms and targets. Therefore, these two error functions have been combined as $e_p(f) - e_c(f)$, such that the system's effectiveness can be assessed using the error ratio.

$$\varepsilon(f) = 10\log_{10}\frac{||\mathbf{d}(f)-\mathbf{c}(f)||^2}{||\mathbf{d}(f)-\mathbf{p}(f)||^2}. \quad (2.32)$$

Here, ε is the error ratio. A plot of either the error function or error ratio shows the same data in a different format, so the error ratio will be plotted against frequency when it is required to ensure comparability of different plots. The error ratio can also be expressed as the mean value $\bar{\varepsilon}_{(f_1-f_2)}$, giving a simple single value that can be compared across all systems. $(f_1 - f_2)$ denotes a banded frequency for which the error is analysed from f_1 to f_2 Hz. In simple terms:

- If $\bar{\varepsilon}_{(f_1-f_2)} < 0$, the control is more effective than if there were no control.
- If $\bar{\varepsilon}_{(f_1-f_2)} = 0$, the control is no more effective than if there were no control.
- If $\bar{\varepsilon}_{(f_1-f_2)} > 0$, the control is less effective than if there were no control.

2.4 Room dimension measurement and mapping

In Chapter 5, the aim is to derive the dimensions of the environment and the location of the sources and the receiver within it to model the space mathematically. The previous methods for achieving this goal are now reviewed.

2.4.1 Enclosure dimensions estimation

Previous work was primarily based on the time domain analysis, including Aprea et al. (2009), Filos et al. (2010), Antonacci et al. (2012), Tervo and Tossavainen (2012), Dokmanić et al. (2013), Remaggi et al. (2014), Filos et al. (2011), Moore et al. (2013), Crocco et al. (2016) and Devos (2019). These methods generally analyse sound intensity vectors and time and direction of arrival (Tervo et al. (2009)) to estimate the distance from the receiver to the surface. Unlike the method presented in Chapter 5, the methods in this section add complexity to the measurement process by either requiring a specific audio signal for the measurement or using microphone arrays. A summary of a number of these methods follows:

1. Antonacci et al. (2012) focuses on two-dimensional geometries using a single source and an array of microphones. This method requires that the relative location of each microphone in the array is known. The system estimates the distance between the surface and the microphone array by comparing the arrival time of the direct signal from the source and the reflected signal from the surface at the location of each microphone in the array.

2. Tervo and Tossavainen (2012) estimates three-dimensional geometries using a single source and a microphone array. The proposed method does not require prior knowledge of the number of walls or the room shape. The method looks at the peak distribution in the time domain impulse response while adjusting the directivity of the microphone array. This allows it to calculate the direction of the surface (as seen by the microphone array) by analysing the energy in the impulse and then the distance to the surface. Although the paper states that the algorithm does not require any a priori information about the room shape, the experimental data is only for a cuboid room, which does not support this claim.

3. Dokmanić et al. (2013) presents an algorithm for reconstructing three-dimensional geometries of a convex polyhedral room from a few acoustic measurements. It uses a single source and recommends using four separate microphones for three-dimensional estimations. The paper also proposes using an impulsive sound stimulus, which is questionable because the study uses a kick drum as the stimulus. However, through an analysis of the direct sound and echoes, by grouping the echoes by the surface they were emitted from, the distance from that surface can be estimated, from which the geometry of the enclosure can then be inferred.

The method introduced in Chapter 5 works in the frequency domain using the data already present from the room impulse response. There is no requirement for a specific test signal or a microphone array, as it can use microphones built into each source.

2.4.2 Location estimation within an enclosure

This section reviews methods for estimating the locations of an acoustic source and receiver within an enclosure. All previous work on this was primarily based on the time domain analysis. There is a large amount of work done in this area using microphone arrays and methods such as beamforming to locate the audio source, including but not limited to that by Weng and Guentchev (2001), Ma and Liu (2018), Rui et al. (2005), Wu et al. (2010), Bodley (2003), Ma et al. (2019), Zhang et al. (2008) and Ribeiro et al. (2010). The main focus of these papers is speech location for voice commutation systems. The work in this thesis aims to use a single microphone built into each source rather than a microphone array, as there is an imperative to keep the hardware cost to a minimum if the approach introduced is going to be transferable to a consumer product. Therefore, beamforming is not an option.

One option is to use a point-to-point measurement system to estimate the distance between the source and receiver, allowing the relative locations to be triangulated once enough data is estimated. Methods such as that proposed by Georganti et al. (2011) measure the source-to-receiver distance through feature extraction from analysing the recording. The more straightforward approach is to measure the time of flight between the source and receiver (Bell (2011)). This thesis aims to use a single microphone built into each source. Accordingly, a time-of-flight system represents a possible solution, and a preliminary study to investigate it is given here.

The time-of-flight method involves playing a signal from the source loudspeaker at point **s** and recording the time delay required to reach the receiver at point **r**, giving a time of flight $sr(t)$ by which the distance $sr(m)$ can be calculated. Once three (or more) loudspeakers are in use, the matrix **SR** will be created, and the distance between the source and the relative locations can be calculated (Bell (2011)).

FIGURE 2.8: Representation of the system setup used for these tests.

Figure 2.8 shows a representation of the system setup. Each loudspeaker had a built-in microphone. The aim was to demonstrate that a real-time impulse-based time of flight system could be used to measure the distance from the source to the receiver in a multi-location system.

TABLE 2.8: Physical distance measurement

		Source			
		1	2	3	4
Receiver	1		2.47	4.42	4.73
	2	2.46		4.71	3.91
	3	4.35	4.66		1.65
	4	4.71	3.88	1.70	

TABLE 2.9: Distance measured with audio impulse

		Source			
		1	2	3	4
Receiver	1		2.44	4.40	4.71
	2	2.46		4.68	3.89
	3	4.34	4.65		1.67
	4	4.70	3.87	1.69	

TABLE 2.10: Error between physical measurements and distance measured with audio impulse

		Source			
		1	2	3	4
Receiver	1		0.03	0.02	0.00
	2	0.00		0.03	0.02
	3	0.01	0.01		0.02
	4	0.01	0.01	0.01	

Table 2.8 shows the physical measurement from each loudspeaker to each receiver, and Table 2.9 shows the measurement from each source to each receiver using the real-time impulse-based time-of-flight system. Table 2.10 shows the error between the two measurements. The largest absolute error in Table 2.10 was 0.03 m, which is 0.6%, and the largest percentage error was 1.2%. Although this worked well, this point-to-point measurement system only provided relative locations. The method in Chapter 5 aims to estimate the absolute locations in a coordinate system consistent with the room.

2.4.3 Other room measurement systems

Several commercial applications have been released to determine the location of a portable device (e.g. a mobile phone or smart loudspeaker) relative to a particular room or space. These location-based services serve different purposes but usually employ radio-based technologies (Tarzia et al. (2011)). The most common approaches for indoor localisation require the following:

- The installation of different sensors within the environment for location accuracy and reliability.

- The use of existing cellular or Wi-Fi base stations. The most accurate Wi-Fi localisation approach uses a Gaussian distribution to statistically model signal strength observations.

An acoustic space analysis has been proposed to overcome the indoor localisation issues associated with radio-based systems. Some of the most interesting technologies and projects are summarised in this section. Again, the methods in this section add complexity to the measurement process, unlike those presented in Chapter 5.

2.4.3.1 Human-generated sound location systems

Human-generated sound location systems require the end-user to generate a sound, which is then used to locate the user or calculate the room impulse response. Examples include:

1. Accurate, low-cost location sensing (Scott and Dragovic (2005)) requiring the placement of microphones in key locations. Users can then snap their fingers to share their positions in space. At least four microphones are required to calculate the sound's x, y and z coordinates.
2. Use of the hand clap as an impulse source for measuring room acoustic (Seetharaman and Tarzia (2012)).

Both these systems work but require a clear signal-to-noise ratio. To give effective results, the hand clap excitation must be 26.4 dB above the background noise level (Seetharaman and Tarzia (2012)). Location sensing requires knowledge of other locations within the environment (where the microphones are placed). Overall, this could offer an effective way to locate a listener within an already-setup system.

2.4.3.2 Multi-sensor systems

These systems use a combination of information from different sensors to locate the listener in the space. The most well-known are the Apple Homepod (Hardy et al. (2018)) and built-in loudspeakers. These systems use built-in video cameras, compasses, accelerometers, light sensors, wireless antennae, thermometers, current and voltage monitors, microphones, gyroscopes and barometers to determine as much as possible about the environment.

The goal of the Homepod (Hardy et al. (2018)) is to identify the loudspeaker location and orientation relative to a target (i.e. the listener) and then to obtain the characteristics of the surrounding environment. To detect the latter, a set of audible and inaudible

test tones are played through all or some of the loudspeakers (depending on the measurement being made). Simultaneously, the microphone array uses different weighting and delays to produce different polar patterns to detect sounds from the loudspeakers and analyse the presence, level and delay between reflections. For example, significant reflections with minimal delay may indicate that one loudspeaker is close to a hard surface.

2.4.4 Section summary

This section reviewed methods that aim to derive the dimensions of an enclosure and the location of a source and receiver within it from acoustic measurements. All of the methods reviewed add complexity to the measurement process. In Chapter 5, the room dimension and source and receiver location estimation methods differ from any of the systems identified. They require no more complexity than that resulting from measuring a room's impulse response using techniques such as that outlined by Farina (2000).

2.5 Chapter summary

This chapter reviewed the previous research necessary to understand this thesis, placing the present work in its proper context.

It began with a review of the relevant areas of room acoustics and an introduction of Green's functions as an analytical method to model measurements within a cuboid room environment. Section 2.1 also discussed how the dimensions and damping of an enclosed listening environment affect the audio, which can negatively affect the audio experience. Section 2.2 then discussed methods for controlling the acoustic fields within enclosures. All the methods reviewed attempted to compensate for the negative effect on the audio experience, using either digital room correction or acoustic cancellation. The extent to which the plane wave cancellation method can objectively and subjectively improve the audio experience was demonstrated, while limitations of this method were also introduced. This led to the aim of this thesis: to investigate whether the least-squares approach to digital room correction for controlling the acoustic field inside enclosures is objectively as good as the plane wave cancellation method while also avoiding its limitations. To fully investigate this, Chapter 3 will further study the plane wave cancellation method, discussing the theories underpinning it and introducing studies that aim to discover more about the method's limitations. Chapter 4 will then directly compare the least-squares approach to digital design filters for digital room correction to the plane wave cancellation method. Section 2.3 has introduced the metrics that will be used throughout the following chapters when comparing the different approaches for controlling the in-room sound field.

The final section of this chapter reviewed methods to derive the dimensions of an enclosure and the location of a source and receiver within the said enclosure from acoustic measurements. Although the methods are effective, they add complexity to the measurement process. This provides the opportunity for further research to determine whether a method can be created that derives the dimensions of an enclosure and the location of the source and receiver without the need for added complexity. This is the focus of Chapter 5.

Chapter 3

An assessment of the plane wave cancellation method

The plane wave cancellation method was introduced in Chapter 2, and when reviewing the work of Celestinos and Nielsen (2004), Celestinos and Nielsen (2005), Celestinos and Nielsen (2006), Celestinos and Nielsen (2008a), Celestinos and Nielsen (2008b), Santillán (2001), Santillán et al. (2007) and Fazenda et al. (2012) in conjunction with the other methods of control outlined in Chapter 2, it was decided that the plane wave cancellation method would be the reference for this work.

This chapter is split into three sections to provide a greater understanding and assessment of the plane wave cancellation method, as follows:

1. A review of theories underpinning the plane wave cancellation method
2. A demonstration of the effectiveness of the plane wave cancellation method
3. An investigation into the limitations of the plane wave cancellation method

After these matters have been discussed, it will be clear how effective the plane wave cancellation method can be and how the method is limited when moving to a home environment, thus creating an opportunity for the use of the approach assessed in Chapter 4.

The results analysed from the studies presented throughout this chapter will look for trends rather than going into too much detail, as this is a short review; more in-depth studies can be found in the cited literature.

3.1 Theories underpinning the plane wave cancellation method

This section introduces and explains theories that create the conditions from which the plane wave cancellation method arises. Two theories will be covered. The first, the principle of superposition, explains why wave cancellation occurs. The second shows how the combination of wave field synthesis (WFS) and the image method can create a plane wave in the real world.

3.1.1 Wave cancellation

The idea behind wave cancellation comes from the principle of superposition: the pressure at any location within a sound field can be expressed as the sum of all sources at that location (Nelson and Elliott (1992)). This can be mathematically shown by starting with the linearised version of the three-dimensional wave equation,

$$\nabla^2 p - \frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2} = 0, \tag{3.1}$$

and substituting in two independent acoustic pressure functions, $p_1(\mathbf{x}_1, t)$ and $p_2(\mathbf{x}_2, t)$ (Nelson and Elliott (1992)), which leads to

$$\left[\nabla^2 - \frac{1}{c_0^2}\frac{\partial^2}{\partial t^2}\right] p_1(\mathrm{x}_1, t) = 0 \qquad \& \qquad \left[\nabla^2 - \frac{1}{c_0^2}\frac{\partial^2}{\partial t^2}\right] p_2(\mathbf{x}_2, t) = 0 \tag{3.2}$$

and therefore

$$\left[\nabla^2 - \frac{1}{c_0^2}\frac{\partial^2}{\partial t^2}\right] (p_1(\mathbf{x}_1, t) + p_2(\mathbf{x}_2, t)) = 0. \tag{3.3}$$

This leads to

$$p_{total}(\mathbf{x}, t) = p_1(\mathbf{x}_1, t) + p_2(\mathbf{x}_2, t). \tag{3.4}$$

From this, it is clear that at the point where two pressure waves meet, if $p_1(\mathbf{x}_1, t) = -p_2(\mathbf{x}_2, t)$, then $p_{total}(\mathbf{x}, t) = 0$, so there will be full cancellation. This mathematical proof can be extended to any number of independent acoustic pressure functions, and this wave cancellation naturally occurs when pressure waves interact. The key to any form of acoustic control that relies on wave cancellation is manipulating how and where these interactions occur.

3.1.2 Why a plane wave?

In audio reproduction systems, most real-world sound sources radiate low-frequency pressure waves in a pattern similar to a sphere. Therefore, this wave propagation is normally simplified as a spherical wave. As shown mathematically in Section 3.1.1, wave

cancellation can occur naturally when pressure waves interact. Any acoustic control system dependent on wave cancellation to change the sound field must predict and manipulate these interactions, either in the environment as a whole or across defined areas.

Figure 3.1 shows how two spherical waves interact in a free-field environment where the wave sources are identical except for a 180° phase difference.

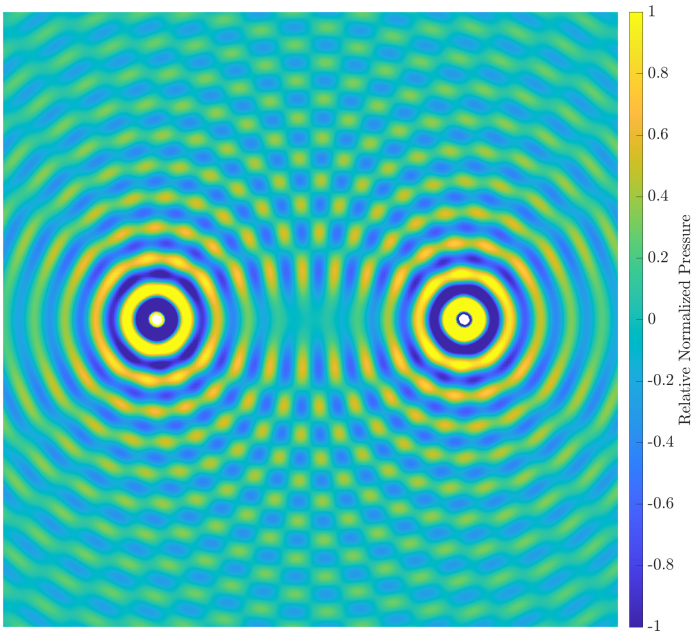


FIGURE 3.1: Spherical wave interaction in the free field. Both sources are identical, except for a phase difference of 180°.

Figure 3.1 shows that although areas of local cancellation (zero pressure) can be observed, with the clearest running directly between the two points; there are also areas of constructive inference. In the free-field environment, these areas of cancellation and reinforcement are distinct. Figure 3.2 shows the same system placed in a simple enclosed environment, where the reflections complicate the wave interactions. The interactions are still predictable in the real world, but they become more complex; thus, it becomes computationally intensive to calculate them. Additionally, as discussed in Chapter 2 when reviewing the work in Nelson and Elliott (1992), Vanderkooy (2007), Vanderkooy (2011) and Vanderkooy and Rousseau (2013), cancellation systems that use spherical waves perform poorly. What is needed is an easy-to-predict, uniform wavefront, which is offered by a plane wave.

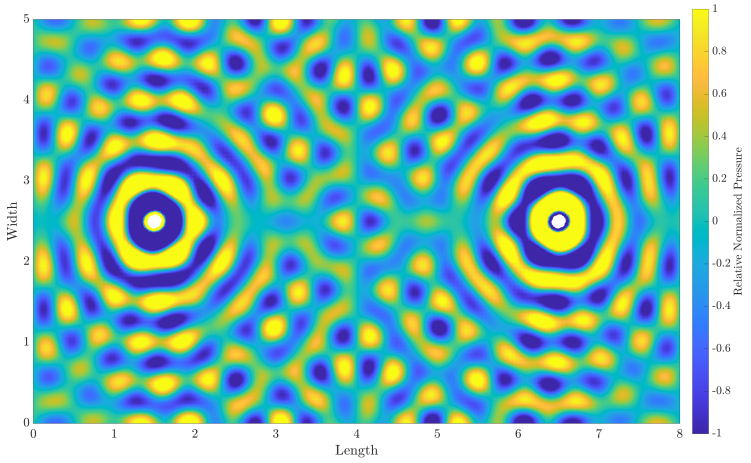


FIGURE 3.2: Spherical wave interaction in an enclosure. Both sources are identical, except for a phase difference of 180°.

3.1.3 Creating a plane wave

If the spherical wave propagation shown in Figure 3.2 can be replaced by a system that can generate two plane waves at either end of the environment, the calculations shown in Section 3.1.1 demonstrate that it should be possible for one wave to be completely cancelled by the second. This is the main concept in Celestinos and Nielsen (2008b). Although an infinite plane wave is not possible from a real-world sound source, it is theoretically possible to create a pseudo-infinite plane wave using many spherical sources.

The procedure that allows any wavefront to be constructed using a number of spherical sources was first developed by Berkhout (1988). This paper led to the theory of WFS, as outlined in Berkhout et al. (1993). Wave field synthesis is a sound field control method designed specifically for spatial audio rendering. Utilising WFS concepts, it should be possible to create a plane wave by placing a number of spherical sources in a line array (Ward and Abhayapala (2001)).

Figure 3.3 shows the horizontal cross-section of 100 free-field spherical sources in a line array. A plane wave can be seen forming in the middle of the array. The next problem is that this number of sources cannot realistically be used in a home environment. However, a virtual array of reflected sources can be created using the image method (Allen and Berkley (1979)). For this, the sidewalls are modelled as flat, pure, reflective surfaces.

Figure 3.4 shows the main room with the real and image sources created by reflecting walls. If the sidewalls are modelled as flat, pure, reflective surfaces, these virtual rooms will extend to infinity on either side, creating an infinite array of sources. In the real world, absorption would occur, but this method could create a close approximation to

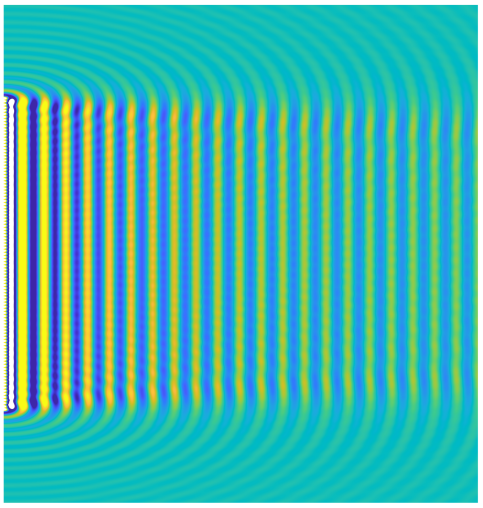


FIGURE 3.3: Use of 100 spherical sources to create a plane wave.

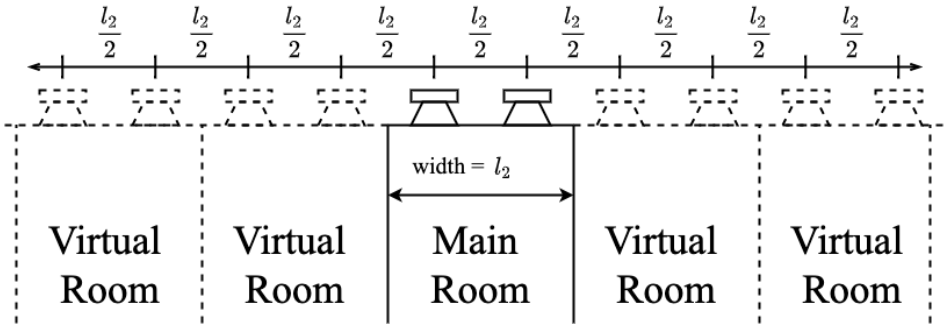


FIGURE 3.4: Creation of an array using an image source.

a plane wave across most of a real room. Figure 3.4 shows the layout for an array with two real-world loudspeakers, but any number of loudspeakers could be used. Their location within the real room is important; when they are mirrored, they must create a uniform array. With W representing a generic number of loudspeakers (or sources) in two dimensions, the distances required between the loudspeakers can be expressed as

$$\Delta_y = \frac{l_2}{W}, \tag{3.5}$$

where Δ_y is the distance in metres between the point sources. The aim is to create the highest-frequency plane wave possible to propagate along the longest dimension of the enclosed space. The limiting factors of the highest frequency at which a plane wave can be created are the room's width and the number of sources. This is known as the aliasing frequency (Spors and Ahrens (2009)), which is given by

$$f_{aliasing} = \frac{c_0}{2\Delta_y}. \tag{3.6}$$

From this, it can be stated that the distance between the loudspeakers fixes the highest frequency at which a plane wave can be created. Combining Equation 3.6 and Equation 3.5, it can be shown that the plane wave cancellation method can only take place for modes up to

$$f = \frac{c_0 W}{2l_2}. \tag{3.7}$$

Using the information given in this section, it should now be possible to create a sound field up to the maximum plane wave frequency, which propagates from both the front and rear of a given cuboid environment and uses the rear waves to cancel the front waves. This removes modal resonances and the reflections from the rear wall and thus gives the impression of an environment with no rear wall.

This thesis focuses on a plane wave in two dimensions created across the enclosure's l_2 dimension. Image sources are also created by reflections in the l_3 dimension (from the ceiling and floor) following the same rules. The upper frequency at which a three-dimensional plane wave can be created is limited by the largest distance between two sources (or one source and an image source) in either the l_2 or l_3 dimension. Focus is on a plane wave in two dimensions, as it has been surmised that most home listening environments have two front sources.

3.2 Testing the plane wave cancellation method

Building on the principles outlined in the previous section, the efficacy of the plane wave cancellation method was assessed through three studies carried out using an analytical model. These studies were less comprehensive than those conducted by Celestinos and Nielsen (2006), Celestinos and Nielsen (2008a), Celestinos and Nielsen (2008b) and Santillán et al. (2007), showing only small sets of data comparing the effectiveness of systems with and without the plane wave cancellation method. However, these three studies also build a benchmark for work done later in this thesis. They covered the following topics:

Study 1: The effect of the plane wave cancellation method on the modal resonances throughout the environment

Study 2: The effect of the plane wave cancellation method at a single point

Study 3: The impact of the plane wave cancellation method across a frequency band over a fixed listening area within the enclosure

The main parameters of each model, namely the cuboid enclosure and the transfer function calculation methods, were identical. The main differences concerned the area within the studied enclosure and how the data was analysed.

3.2.1 Shared parameters used in the studies

There were two sets of shared parameters used in these studies, as follows:

1. The dimensions of the enclosure and the locations of the sources within the enclosure
2. The mathematical model used to create the transfer functions

3.2.1.1 The enclosure and the sources' locations

The modelled room was based on the Bowers & Wilkins SRE (Steyning research establishment) main listening room. The dimensions (L) = ([length, width and height]) were as follows: $(\mathrm{l}_1, \mathrm{l}_2, \mathrm{l}_3) = (\,8\,\mathrm{m},\ 4.93\,\mathrm{m},\ 2.96\,\mathrm{m})$. Any point within the room was defined in three-dimensional Cartesian coordinates in the form of $(x\,,\ y,\ z)$, with the original reference at the front-bottom left-hand corner (unless otherwise stated).

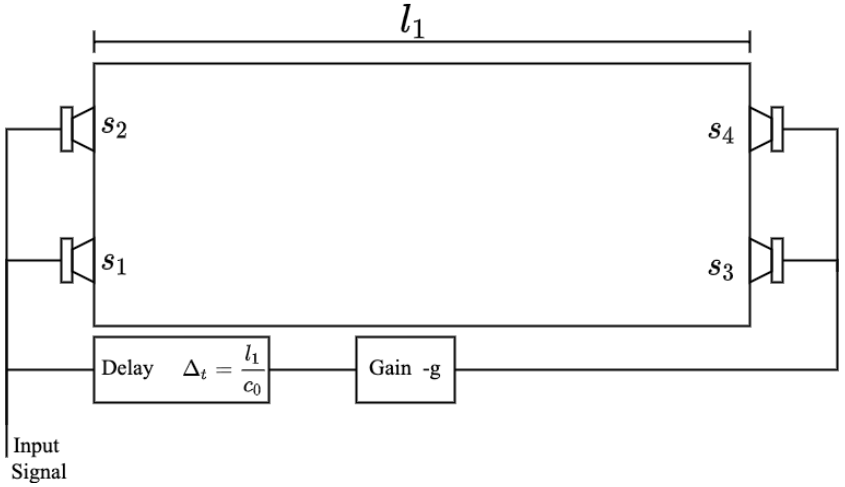


FIGURE 3.5: A basic plane wave cancellation method room representation shown from above.

In Figure 3.5, a representation of the loudspeaker layout is shown. Throughout this thesis, the signal sources are defined as s_w with a location of $\mathbf{s}_w = (\tilde{x}_w, \tilde{y}_w, \tilde{z}_w)$. The source locations for the studies in this section are given in Table 3.1. Following the theory outlined in Section 3.1.3, these locations are the ideal source positions to create

plane wave propagation along the l_1 dimension at frequencies up to 69.57 Hz from both pairs of sources.

TABLE 3.1: The ideal source positions for the plane wave cancellation method. The rear channels ($\mathbf{s}_3$ and $\mathbf{s}_4$) are modelled for all plane wave cancellation methods at $\tilde{x}$ values $\mathbf{s}_3$ and $\mathbf{s}_4$ as $2l_1$

Source No.	Source Location ($\tilde{x}$ m, $\tilde{y}$ m, $\tilde{z}$ m)	Source No.	Source Location ($\tilde{x}$ m, $\tilde{y}$ m, $\tilde{z}$ m)
$\mathbf{s}_1$	(0.0, 1.23, 1.48)	$\mathbf{s}_3$	(8.0, 1.23, 1.48)
$\mathbf{s}_2$	(0.0, 3.70, 1.48)	$\mathbf{s}_4$	(8.0, 3.70, 1.48)

Figure 3.5 shows the method used to achieve plane wave cancellation as given from the CABS set up in Celestinos and Nielsen (2008b), where the rear channels transmitted the same signal as the front channels, delayed by an amount proportional to the room's length, which was then inverted by adding a negative gain. This gain was proportional to the level of signal attenuation between the front and rear walls (although the authors do not provide the gain factor or method of calculation in their paper). A version of this method has been used for plane wave cancellation throughout this thesis. However, because the plane wave cancellation method is only tested in a modelled environment, the requirement to delay, attenuate and invert the phase of the rear channels ($\mathbf{s}_3$ and $\mathbf{s}_4$) can be negated by simply setting the $\tilde{x}$ values of $\mathbf{s}_3$ and $\mathbf{s}_4$ to $2l_1$. This automatically creates the most effective plane wave cancellation possible in the space, by creating a second identical modelled room behind the original. As these two rooms are acoustically identical, the signals leave the front sources $\mathbf{s}_1$ and $\mathbf{s}_2$ and the rear sources $\mathbf{s}_3$ and $\mathbf{s}_4$ at the same time, and are attenuated by the same amount as they travel through their respective rooms until they interact at the adjoining surface. Here the signal from the rear sources will be a perfect inverse of the signal from the front sources.

3.2.1.2 Transfer function

The transfer function for the analytical model was created using the Green's function pressure model, previously shown in Equation 2.23, as follows:

$$\hat{\mathrm{P}}(\mathbf{r}|\mathbf{s}) = \sum_{n=0}^{\infty} \frac{\omega\rho_0 c_0^2 Q\ \psi_n(\mathbf{r})\psi_n(\mathbf{s})}{[2\zeta_n\omega_n\omega + i(\omega^2 - \omega_{\mathbf{n}}^2)]\mathrm{V}} \tag{3.8}$$

It gives the radiated pressure at a point, $\mathbf{r}_m$, due to a source, $\mathbf{s}_w$, where Q is the volume velocity (in the scope of this thesis, it can be noted that $Q = 1$ unless otherwise stated) and $\zeta_{\mathbf{n}}$ is the damping coefficient. This model was originally designed as a

pseudo-approximation of a real-world environment, although with greatly simplified aspects, such as a lack of objects within the space, a lack of windows or doors and all surfaces being completely smooth with uniform damping. However, the aim was that the damping coefficient should represent one that might be found in the real world. Therefore, in Section 2.1.1, it was stated from Hoeg et al. (1997) and the technical document AES (2001), that the target RT_{60} should be between 0.2 and 0.5 seconds in the frequency range from 250 Hz to 4 kHz, with Beranek and Mellow (2019) revealing that at low frequencies, this target may rise 1.25 %. In Section 3.1.3, it was discussed that the lower the absorption of the surface, the more effective the image method is at constructing a plane wave. To provide the highest reflective surface based on real-world average targets, an RT_{60} of 0.625 seconds was used for the studies in this section. Utilising Equation 2.25 from Section 2.1.3.3, this RT_{60} value gives a $\zeta_{\mathbf{n}}$ of $\frac{11.052}{\omega_n}$.

The accuracy of the analytical model needs to be discussed. In a room where the l_1 dimension is 8 m and a speed of sound of 343 ms^{-1}, the first resonance frequency would be 21.4375 Hz. However, due to the computational effort required, the Green's function model is limited to a frequency resolution of two decimal places. The offset and error created in the cancellation are deemed small enough not to affect the broader outcomes of the work, as any realistic, usable level of improvement must be much larger than this tolerance.

In each study, the raw signal without cancellation is compared to the plane wave cancellation method signal. This raw signal is the summation of the transfer function from s_1 and s_2 and is the same throughout the thesis unless otherwise stated.

3.2.2 The plane wave cancellation method whole-room study

This study looked at the effect of the plane wave cancellation method on individual modal resonances throughout the whole environment. The modelled room was sampled with a grid with a receiver point at every 0.10 m in the x direction and 0.0986 m in the y. This grid was a flat surface at a height of 1.48 m. The transfer function, as defined in Equation 3.8, was then used to calculate the frequency response from each source to each receiver point. These frequency responses were then combined so that any individual frequency could be analysed across the whole surface.

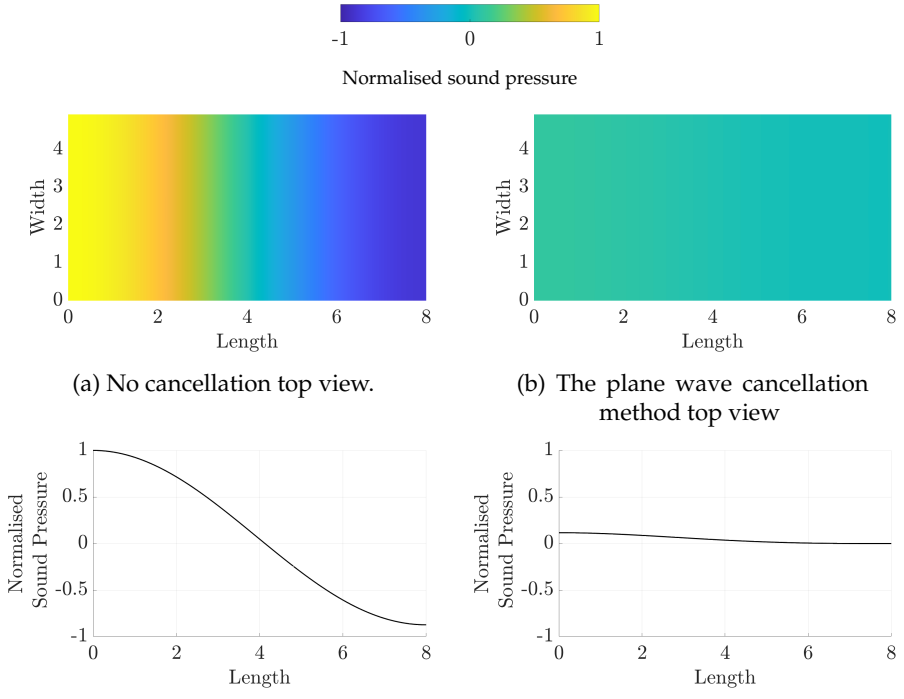
This section reviews the results for the first three $(n_1, 0, 0)$ modes. In Table 3.2, the first 20 modes of the enclosure are reported, calculated using Equation 2.9, and the first three $(n_1, 0, 0)$ modes' frequencies were 21.44 Hz, 42.88 Hz and 64.31 Hz.

Figures 3.6, 3.7 and 3.8 show the plots for the sound pressure level over the defined surface grid across the whole room for modes 21.44 Hz $(1, 0, 0)$, 42.88 Hz $(2, 0, 0)$ and 64.31 Hz $(3, 0, 0)$, respectively. Each figure is made up of four plots: (a) the acoustic pressure on the surface with no cancellation, (b) the acoustic pressure on the surface with the

TABLE 3.2: First 20 modal frequencies based on the dimensions of the Bowers & Wilkins SRE main listening room.

(n_1, n_2, n_3)	Frequency	(n_1, n_2, n_3)	Frequency	(n_1, n_2, n_3)	Frequency	(n_1, n_2, n_3)	Frequency
(0,0,0)	0.00 Hz	(2,1,0)	55.21 Hz	(0,2,0)	69.57 Hz	(2,1,1)	80.03 Hz
(1,0,0)	21.44 Hz	(0,0,1)	57.94 Hz	(1,1,1)	70.90 Hz	(2,2,0)	81.72 Hz
(0,1,0)	34.79 Hz	(1,0,1)	61.78 Hz	(2,0,1)	72.08 Hz	(4,0,0)	85.75 Hz
(1,1,0)	40.86 Hz	(3,0,0)	64.31 Hz	(1,2,0)	72.80 Hz	(3,0,1)	86.56 Hz
(2,0,0)	42.88 Hz	(0,1,1)	67.58 Hz	(3,1,0)	73.12 Hz	(0,2,1)	90.54 Hz

plane wave cancellation method, (c) the sound pressure as a function of the x coordinate with no cancellation and (d) the sound pressure as a function of the x coordinate with the plane wave cancellation method. The plots within each figure have been normalised by the maximum sound pressure value at the frequency under analysis, with no cancellation.



(a) No cancellation top view.

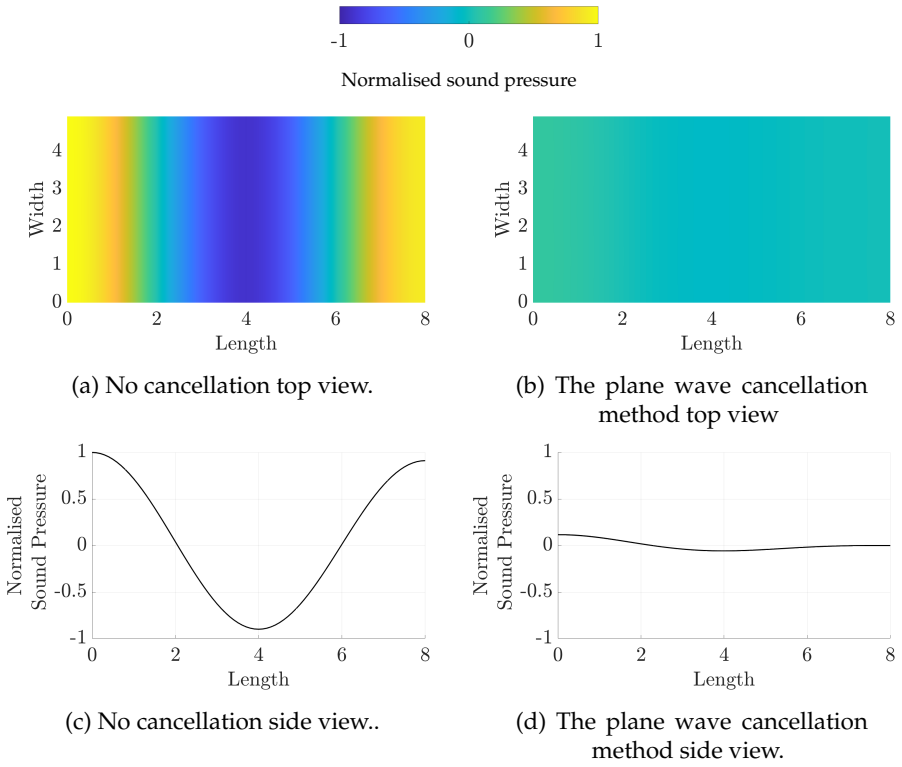
(b) The plane wave cancellation method top view

(c) No cancellation side view..

(d) The plane wave cancellation method side view.

FIGURE 3.6: The sound pressure in the enclosed environment at 21.44 Hz, with and without the plane wave cancellation method.

The plots in Figures 3.6, 3.7 and 3.8 all show a drop in the relative normalised sound pressure when the plane wave cancellation method was in use. The drop is clearest in the difference between the (c) and (d) plots in each figure. A comparison between the plots shows how the peak pressure at 0 m decreased by 88.4 % in Figure 3.6, 88.2 % in Figure 3.7 and 88.0 % in Figure 3.8. Table 3.3 shows the drop in the RMS SPL across the whole surface at each frequency as a decibel value.



(a) No cancellation top view.

(b) The plane wave cancellation method top view

(c) No cancellation side view..

(d) The plane wave cancellation method side view.

FIGURE 3.7: The sound pressure in the enclosed environment at 42.88 Hz, with and without the plane wave cancellation method.

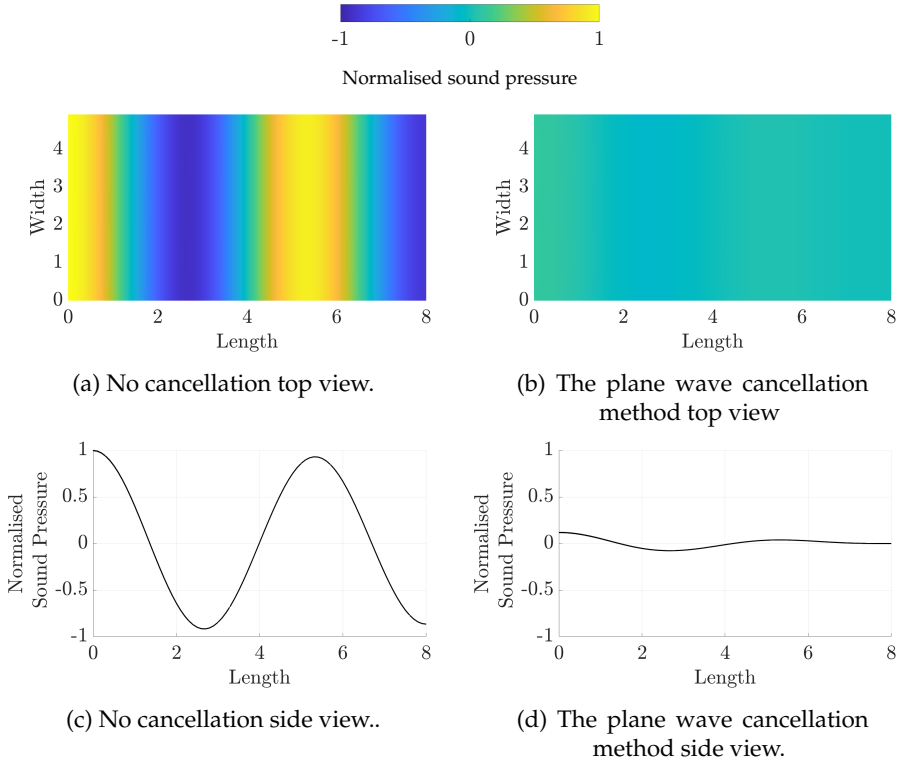
TABLE 3.3: The RMS levels of the raw signal with no cancellation and the plane wave cancellation method signal over the surface at modes 21.44 Hz, 42.88 Hz and 64.31 Hz and the delta.

Frequency	RMS [dB]		
	RS	PWC	Δ
21.44 Hz	68.11	61.43	-6.67
42.88 Hz	68.30	61.35	-6.95
64.31 Hz	68.25	61.10	-7.15

Table 3.3 shows that the model RMS SPL across the whole surface dropped by up to 7.15 dB when using the plane wave cancellation method.

3.2.3 The plane wave cancellation method single-point study

This study looked at the effectiveness of the plane wave cancellation method across a frequency range at a single point. The receiver was modelled in the corner at location $(0, 0, 0)$, where $\psi_{\mathbf{n}}(\mathbf{r}_m) = 1$ for any value of n, so that the receiver location would have



(a) No cancellation top view.

(b) The plane wave cancellation method top view

(c) No cancellation side view..

(d) The plane wave cancellation method side view.

FIGURE 3.8: The sound pressure in the enclosed environment at 64.31 Hz, with and without the plane wave cancellation method.

no effect on the modal resonances found in the frequency response of this model. This means that this location would produce the highest modal resonance influence for this space with this source location. The data sets were analysed in the frequency band of 10−100 Hz, i.e. the band of interest for much of the work in this thesis.

Figure 3.9 shows the frequency response for the raw signal (—RS) with no cancellation and the plane wave cancellation method signal (—PWC) for the receiver at point $(0,0,0)$, with all modal frequencies of the enclosure denoted as - -.

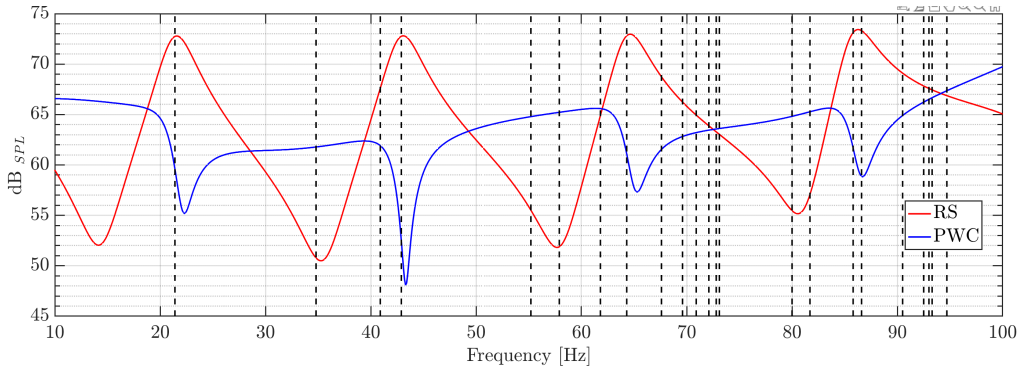


FIGURE 3.9: Frequency response for the RS (—RS) with no cancellation and the PWC (—PWC) for the receiver at point $(0,0,0)$, with all mode frequencies of the enclosure denoted as - -.

In Figure 3.9, the RS shows that only four modal frequencies were excited at the receiver point: 21.44 Hz, 42.88 Hz, 64.31 Hz and 85.76 Hz. The PWC shows high drops in RMS level at these frequencies; the values are given in Table 3.4.

The lowest point in each plane wave cancellation method trough in Figure 3.9 was slightly offset from the modal frequency. This was due to two factors: the model's accuracy, as discussed, and the damping coefficient. As explained in Section 3.1.3, the more image sources there were, the closer the result was to a plane wave. These drops were created when the two waves were not perfect plane waves; as the damping coefficient moved toward zero, the troughs became narrower. Additionally, when the damping coefficient was higher, the model showed lower peaks and had more modal crossover, which also removed the troughs.

In Figure 3.9, the RS also shows that three troughs at 33 Hz, 55 Hz and 80 Hz were improved by 11 dB, 14 dB and 9 dB, respectively, demonstrating that the plane wave cancellation method was correct for both the peak and the local troughs. The full effect of the plane wave cancellation method is shown in the RMS values in Table 3.5.

TABLE 3.4: The levels of the RS with no cancellation and the PWC for the receiver at point (0, 0, 0) at modes 21.44 Hz, 42.88 Hz, 64.31 Hz and 85.76 Hz and the delta.

Frequency	SPL_{RMS} [dB]		
	RS	PWC	Δ
21.44 Hz	72.78	58.27	-14.52
42.88 Hz	72.80	49.77	-23.04
64.31 Hz	72.96	59.50	-13.46
85.76 Hz	73.42	59.51	-13.92

Tables 3.4 and 3.5 show the RMS SPL levels for the RS with no cancellation and the PWC for the receiver at point $(0, 0, 0)$. In Table 3.4 for the data for the modes at 21.44 Hz, 42.88 Hz, 64.31 Hz and 85.76 Hz is given, and in Table 3.5 the data for the whole frequency band of $10-100$ Hz is given. At the single modes, large drops of up to 23.04 dB are seen, and as discussed, the troughs at these frequencies may have been misleading. However, Table 3.5 shows that there is an average drop of 2.52 dB over the frequency range of 10 to 100 Hz at this location when using the plane wave cancellation method.

TABLE 3.5: The RMS levels of the RS with no cancellation and the PWC for the receiver at point (0, 0, 0) for the range of 10 Hz to 100 Hz and the delta.

SPL_{RMS} [dB]		
RS	PWC	Δ
66.02	63.53	2.52

3.2.4 The plane wave cancellation method fixed-listening-area study

This study focused on the impact of the plane wave cancellation method across a fixed listening area within the enclosure. It aimed to provide methods for comparing the plane wave cancellation method with other acoustic control methods throughout this thesis. In this section, the metrics introduced in Section 2.3 will be utilised. Figure 3.10 shows a representation of the listening area, with the receivers defined in terms of r_m with a location of $\mathbf{r}_m = (x_m, y_m, x_m)$. The exact layouts of these receivers are given in Table 3.6.

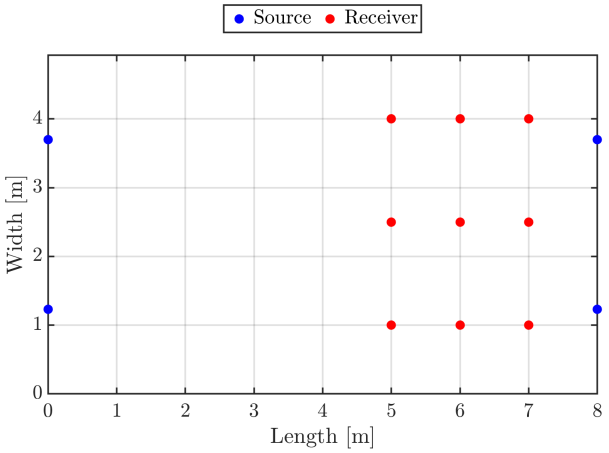


FIGURE 3.10: Representation of the listening area, shown as a grid of receivers denoted as ●, and sources denoted as ●.

TABLE 3.6: Receiver locations.

Receiver No.	Receiver Location (x m, y m, z m)	Receiver No.	Receiver Location (x m, y m, z m)	Receiver No.	Receiver Location (x m, y m, z m)
$\mathbf{r}_1$	(5.0, 1.0, 1.5)	$\mathbf{r}_4$	(6.0, 1.0, 1.5)	$\mathbf{r}_7$	(7.0, 1.01.5)
$\mathbf{r}_2$	(5.0, 3.0, 1.5)	$\mathbf{r}_5$	(6.0, 3.0, 1.5)	$\mathbf{r}_8$	(7.0, 3.0, 1.5)
$\mathbf{r}_3$	(5.0, 4.0, 1.5)	$\mathbf{r}_6$	(6.0, 4.0, 1.5)	r_9	(7.0, 4.0, 1.5)

A target sound field over the listening area was defined as a plane wave travelling through the listening area, expressed as

$$d(x) = p_0 e^{-ik\mathbf{r}_m}. \tag{3.9}$$

The wave-front is propagating parallel to the y-z plane, meaning that $\mathbf{r}_m$ could be simplified to only the x_m value of any receiver point location. In addition, the value p_0 is adjusted so that the target level considers the drop in the average RMS level across the enclosure when the plane wave cancellation method is in use. The goal is to have the plane wave cancellation method show its best results for any setup. In Table 3.5, the

drop in the average RMS level is shown as 2.52 dB for the single point measurement. However, when moving to a fixed listening area, the same raw signal gave an average RMS level of 65.28 dB, and as shown in Table 3.7, the drop in the average RMS level is 1.75 dB when using the plane wave cancellation method. Therefore, throughout this thesis, unless otherwise stated when using the analytical model, the raw signal will have an average RMS level of 65.28 dB over the fixed listening area, with a target level of 63.53 dB. This target level has been chosen as it is the natural level achieved by the plane wave cancellation method. This RMS target level will be added to the metrics introduced in Section 2.3, as it gives a signal which shows how close the system under test is in sound pressure to the target over the target area. This, when included with the other metrics, should build a picture of the effectiveness of the control method under test.

3.2.4.1 Results from the fixed-listening-area study

Figure 3.11 shows the SA for the RS, the PWC and the target, over the listening area. Across the majority of the frequency band, the SA plot shows that the plane wave cancellation method offered a good level of improvement over the RS. However, it shows a number of places where the plane wave cancellation method is further from the target than the RS i.e. the 43 and 85 Hz modes and the section just above 21 Hz. However, the SA plot shows only the average magnitude. This is a reason for plotting the error function or error ratio, as these take both phase and magnitude into account. Figure 3.12 shows how the results of the error function defined in Equation 2.31

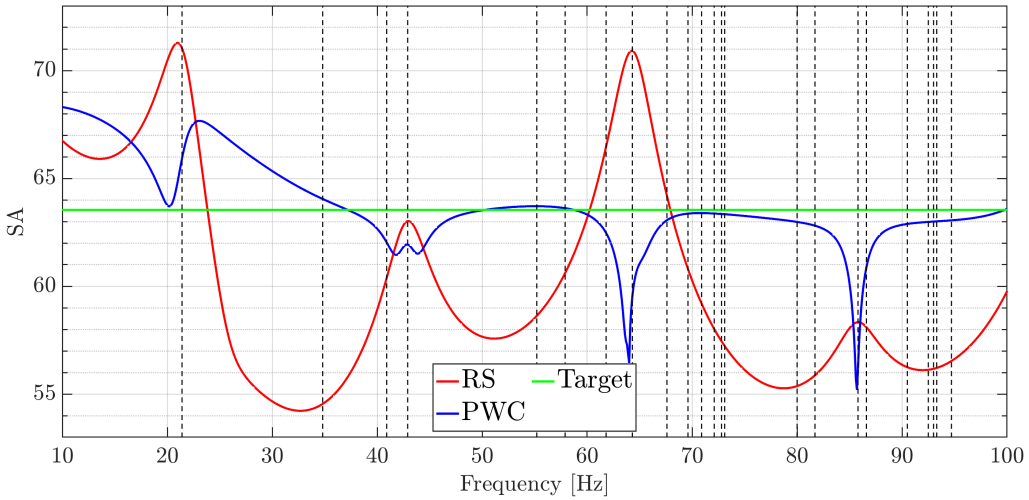


FIGURE 3.11: The SA for the RS (—RS), the PWC (—PWC) and the target (—Target) over a fixed listening area, with all mode frequencies of the enclosure denoted as - -.

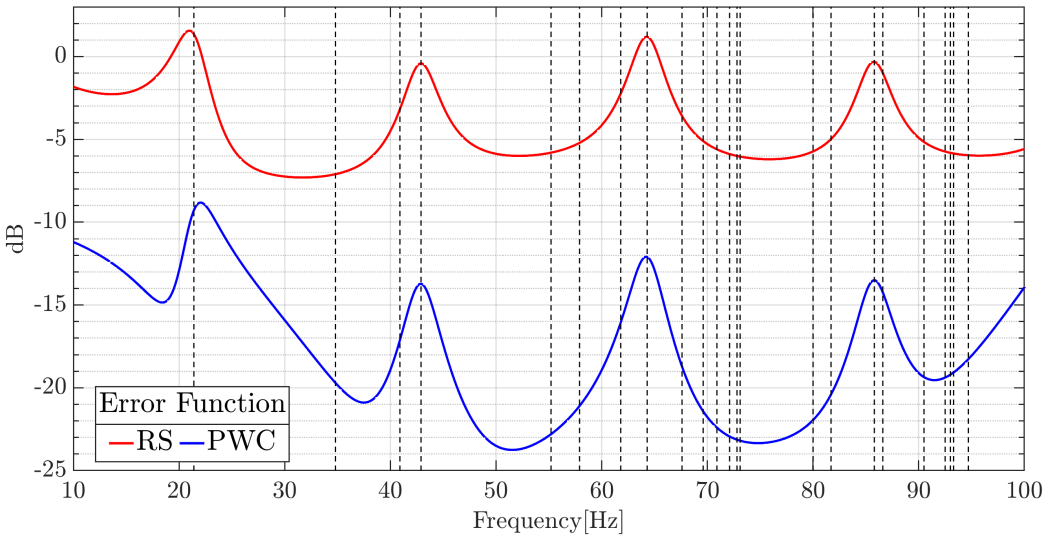


FIGURE 3.12: Error function against frequency for the RS (—RS) with no cancellation and the PWC (—PWC) over a fixed listening area, with all mode frequencies of the enclosure denoted as - -.

The results in Figure 3.12 show that the plane wave cancellation method was closer to the target than the RS. In the error function and error ratio data, the lower the values, the more effective a given method is at achieving the target. The issues highlighted in Figure 3.11 are not seen in Figure 3.12 because the error function took both the amplitude and the phase into consideration. It is interesting that the main modal frequencies 21.44 Hz, 42.88 Hz, 64.31 Hz and 85.76 Hz show peaks in both the RS and PWC data. However, there is an improvement of over 12 dB in the PWC data, showing that although the plane wave cancellation method is moving closer to the target, it still does not offer complete modal control.

Table 3.7 shows the data for the $\bar{\varepsilon}_{(10-100\,\mathrm{Hz})}$, RMS and MSV. These data sets are given in this table format throughout the thesis. The absolute value of the RMS delta is used, as it is deemed more straightforward to cross-compare different targets and approaches, and the distance from each target level is more important than the level itself. Additionally, the data sets shown in Tables 3.7 show the full frequency band of interest and three sub-bands to allow for a greater understanding of the results.

TABLE 3.7: Frequency-banded $\bar{\varepsilon}_{(10-100\,\mathrm{Hz})}$ [dB], RMS [dB], and MSV [dB] over a fixed listening area.

Method	$10\,\mathrm{Hz} > f \geq 100\,\mathrm{Hz}$			$10\,\mathrm{Hz} > f \geq 40\,\mathrm{Hz}$			$40\,\mathrm{Hz} > f \geq 70\,\mathrm{Hz}$			$70\,\mathrm{Hz} > f \geq 100\,\mathrm{Hz}$		
	$\bar{\varepsilon}$	$\lvert\mathrm{RMS}_\Delta\rvert$	MSV	$\bar{\varepsilon}$	$\lvert\mathrm{RMS}_\Delta\rvert$	MSV	$\bar{\varepsilon}$	$\lvert\mathrm{RMS}_\Delta\rvert$	MSV	$\bar{\varepsilon}$	$\lvert\mathrm{RMS}_\Delta\rvert$	MSV
RS	0.00	1.75	46.17	0.00	2.14	23.93	0.00	2.67	47.72	0.00	0.01	67.00
PWC	-12.24	0.00	5.85	-9.80	1.73	2.73	-15.13	1.21	7.85	-13.62	1.26	6.99

In Table 3.7, all of the data shows that the plane wave cancellation method outperformed the RS. Table 3.7 shows that in the 10−100 Hz range, the $\bar{\varepsilon}_{(10-100\,\mathrm{Hz})}$ has a system effectiveness of 12.24 dB, with a maximum effectiveness of 15.13 dB in the sub-range 40−70 Hz and minimum effectiveness of 9.80 dB in the sub-range 10−40 Hz. The RMS target was achieved across the full range (as discussed, this target was chosen for the plane wave cancellation method to give the best results), but there was an improvement greater than 0.41 dB in all sub-ranges. The MSV values in Table 3.7 show that in all but the sub-range 10−40 Hz, the seat-to-seat variation of the RS is above the level of 35 dB (as outlined in Section 2.3 as unacceptable), and even in this sub-range, it is only in the acceptable range.The plane wave cancellation method has reduced the seat-to-seat variation substantially, by 40.17 dB over the full 10−100 Hz range, with all the data below the 15 dB threshold denoting a good level of seat-to-seat variation.

In Chapter 2, the CABS approach for plane wave cancellation was discussed, and the mean sound field deviation metric (MSFD) was introduced from the paper Celestinos and Nielsen (2008a), which is defined as the spatial deviation (SD) and the magnitude deviation (MD).

TABLE 3.8: Comparing the plane wave cancellation method modelled in this thesis with the CABS simulation data taken from Celestinos and Nielsen (2008b).

Setup	MSFD	
	SD	MD
CABS 0.2.0	±4.9	±6.6
CABS 0.2.2	±0.7	±1.9
RS	±5.7	±6.8
PWC	±0.9	±2.0

Table 3.8 compares the MSFD for the plane wave cancellation method modelled in this thesis with the CABS simulation data taken from Celestinos and Nielsen (2008b). It was not expected that this data would be identical, as the room modelled in this thesis has dimensions of 8 m, 4.93 m and 2.96 m. In contrast, the CABS room dimensions were 7.80 m, 4.20 m and 2.76 m, and the CABS receiver array is placed in the centre of the room, whereas in this work the receiver array centred around the location (6.0 m, 3.0 m, 1.5 m). From the data in Table 3.8 it is clear that in absolute terms, the CABS 0.2.2 simulation data gives lower values of both SD and MD than the PWC data, but the delta is small between the CABS 0.2.2 and the PWC 0.2 dB in the SD and 0.1 dB in MD. However, in relative terms, the PWC shows a larger drop from the RS data than the CABS 0.2.2 data from the CABS 0.2.0 data. Therefore, with only a small delta in the absolute values and the larger relative drop from the RS, it can be concluded that the method in this thesis is comparable with the CABS simulation in Celestinos and Nielsen (2008b). This also validates that the model put forward in the chapter can be used throughout this thesis as a reference.

3.2.5 Study review

The primary goal of the studies in this section was to show that the plane wave cancellation method is an effective method of modal control compared with the RS in a cuboid environment. While not going into the level of detail found in the works by Celestinos and Nielsen (2004), Celestinos and Nielsen (2005), Celestinos and Nielsen (2006), Celestinos and Nielsen (2008a), Celestinos and Nielsen (2008b), Santillán (2001) and Santillán et al. (2007), the results shown here still concur with the findings of these papers. From the data in this section, the following can be stated:

- Section 3.2.2 : The RMS level of the first three modes could be reduced by up to 7.15 dB across the whole listening area.
- Section 3.2.3 : The peak level of the first four modes could be reduced by up to 23.04 dB at a single point.
- Section 3.2.3 : The trough levels in the $10-100$ Hz band could be increased by up to 14 dB at a single point.
- Section 3.2.3 : The RMS level in the $10-100$ Hz band could be reduced by up to 2.52 dB at a single point.
- Section 3.2.3 : The $\bar{\varepsilon}_{(10-100\text{ Hz})}$ show a system effectiveness of 12.24 dB over a fixed listening area in the range of $10-100$ Hz.
- Section 3.2.3 : The MSV shows a reduction in the seat-to-seat variation of 40.32 dB over a fixed listening area in the $10-100$ Hz range, bringing the MSV below the 15 dB threshold, which denotes a good level of seat-to-seat variation.
- Section 3.2.3 : The MSFD shows that the plane wave cancellation method model created from this thesis gives a comparable level of effectiveness to the CABS simulation data taken from Celestinos and Nielsen (2008b). This validates that the plane wave cancellation method model created in this chapter gives a good enough result to be used as a reference system throughout the thesis.

3.3 The limitations of the plane wave cancellation method

Section 3.2 shows how effective the plane wave cancellation method system can be when used in an ideal setup. Moving into the home environment, the need for a combined listening and living space will likely mean that the plane wave cancellation method system setup will be compromised. O'Neill (2020) showed that British living rooms are decreasing in size, so there is more competition for floor space with other

furniture. Bell (2011) showed that in the home environment, 5.1 surround sound systems are, on average, not set up as defined in ITU-R:BS.775-3 (2012). Therefore, it is not unreasonable to suggest that, unless the loudspeakers are wall mounted, they may not be in the ideal positions for the plane wave cancellation method system. Additionally, there is no guarantee that the home environment will be precisely cuboid. Adding the need for doors and windows and the fact that walls, floors and ceilings may all comprise different construction materials, the surfaces' acoustic profiles are unlikely to be uniform.

This section will build an understanding of how moving away from an ideal setup could limit the effectiveness of the plane wave cancellation method. It could be argued that it is unfair to test the plane wave cancellation method outside of its fixed parameters. Nevertheless, it is necessary to know the limits of the plane wave cancellation method to create a benchmark for the approach assessed in Chapter 4.

Based on the reliance on the principles of the image method and WFS to create a plane wave, as discussed in Section 3.1, two areas need to be investigated to determine the limitations:

- **Room characteristics:** two main room characteristics affect the ability to build a plane wave: the shape and the reverberation time. The shape will be kept constant. Changing it would bring in too many degrees of freedom, so the focus will be on the reverberation time.
- **Source locations:** to create a plane wave propagation from both ends of a room, an accurate source location is vital.

To test the limitations, three studies were carried out:

Study 1: The effect of damping on the plane wave cancellation method

Study 2: The effect on the plane wave cancellation method of moving s_1 & s_2 perpendicular to the front wall

Study 3: The effect on the plane wave cancellation method of moving s_1 & s_2 parallel to the front wall

All studies in this section used the same room and transfer function as laid out in Section 3.2.1 and the receiver locations defined in Table 3.6. They also used the sources in the locations given in Table 3.1. However, in studies **2** and **3**, moving the locations of s_1 and s_2 from the ideal position was investigated.

3.3.1 The effect of damping on the plane wave cancellation method

In Section 3.2.1.2, the damping coefficient, $\zeta_{\mathbf{n}}$, was set as $\frac{11.052}{\omega_n}$ (RT_{60} = 0.625 seconds) to give the highest reflective surface values while still based on real-world average targets, as outlined by Hoeg et al. (1997), AES (2001) and Beranek and Mellow (2019). Based on the same literature, a real-world minimal RT_{60} would be 0.25 seconds, giving a $\zeta_{\mathbf{n}}$ of $\frac{27.631}{\omega_n}$. These two values were tested here with three linear RT_{60} steps at 0.5312 seconds ($\zeta_{\mathbf{n}} = \frac{13.003}{\omega_n}$), 0.4375 seconds ($\zeta_{\mathbf{n}} = \frac{15.789}{\omega_n}$) and 0.3438 seconds ($\zeta_{\mathbf{n}} = \frac{20.095}{\omega_n}$). In this section, no matter the damping coefficient, the average RMS level of the raw signal was set to 66.02 dB over the listening area, with the average RMS level of the target at 63.53 dB. This was done to ensure no level loss due to the damping, the idea being that a listener in any of the setups would want the raw signal level perceptually to be the same.

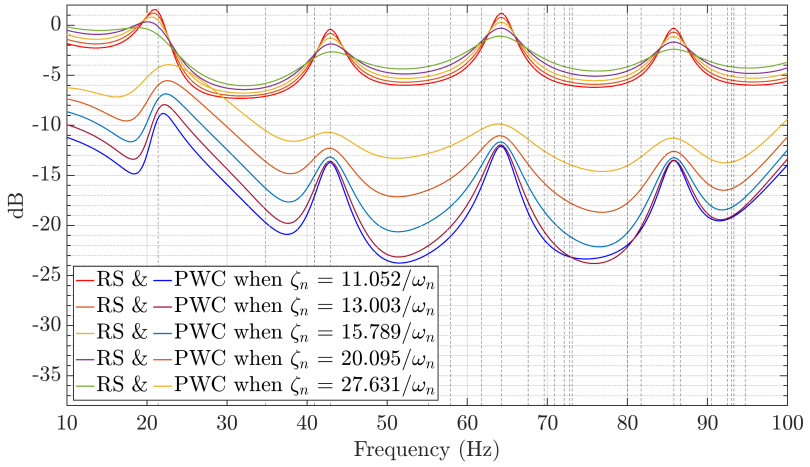


FIGURE 3.13: Error function against frequency for the RS (—RS) with no cancellation and five PWCs with different damping coefficients over a fixed listening area. Original damping value shown as (—), with all modal frequencies of the enclosure denoted as - -.

Figure 3.13 shows how the results of the error function defined in Equation 2.31 changed with the damping. Due to the changes in RS for each value of $\zeta_{\mathbf{n}}$ it is not clear how the effectiveness of each data set compares. However, plotting the error ratio as defined in Equation 2.32 makes the comparison clearer by normalising the RS as shown in Figure 3.14. Throughout the thesis, the error ratio will be plotted rather than the error function when multiple methods are shown on the same plot. Both plots fundamentally show the same thing, but the error ratio is a clearer format for the plot data. The error ratio graphs show the RS at 0 dB (the lower the values for the method under investigation, the more effective the method). Where the values are higher than 0 dB, the method under investigation created a sound field over the listening area that was further from the target than the original RS.

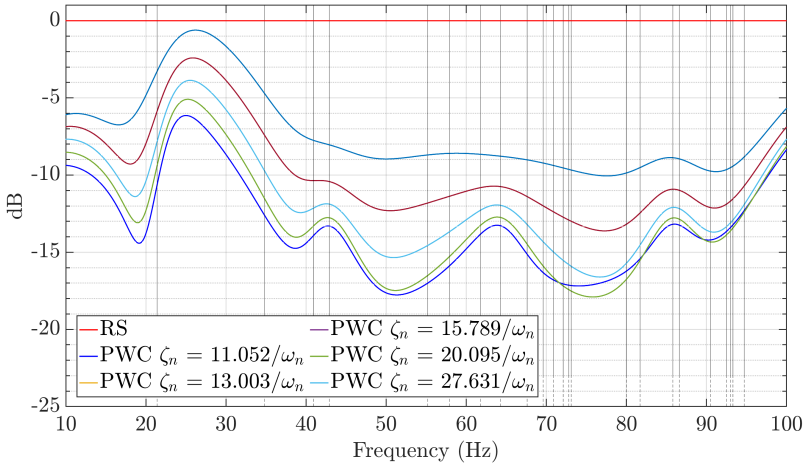
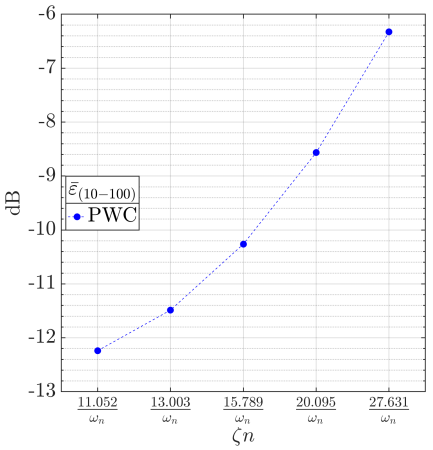


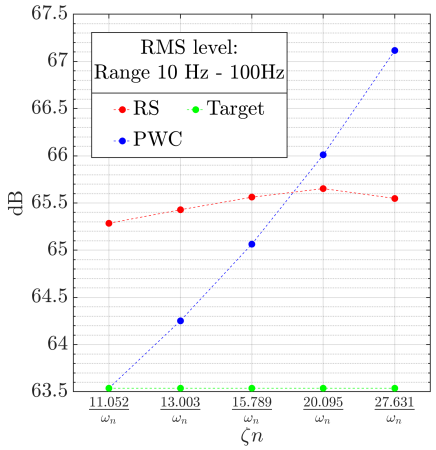
FIGURE 3.14: Error ratio against frequency for the RS (—RS) with no cancellation and five PWCs with different damping coefficients over a fixed listening area. Original damping value shown as (—), with all modal frequencies of the enclosure denoted as - -.

Figure 3.14 shows how the results of the error ratio defined in Equation 2.32 changed with the damping. Only in the band between 20 Hz and 30 Hz was the system with a damping of $\zeta_{\mathbf{n}} = \frac{11.052}{\omega_n}$ outperformed by the other systems; over the rest of the frequency range, the lower the damping value, the higher was the effectiveness observed. This was reinforced in Plot 3.15a, where the $\bar{\varepsilon}_{(10-100\,\mathrm{Hz})}$ averaged over the frequency range clearly shows that as the damping increased, the effectiveness decreased.

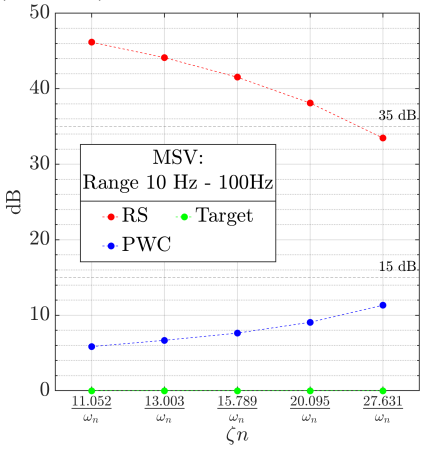
Figure 3.15 shows the error $\bar{\varepsilon}_{(10-100\,\mathrm{Hz})}$ averaged over frequency (Plot 3.15a), the change in the average RMS level (Plot 3.15b), and the MSV over the frequency range in Plot 3.15c as the damping changes. When $\zeta_{\mathbf{n}}=\frac{11.052}{\omega_n}$, it was expected that the RMS level would match the target, as shown and discussed in Section 3.2. In Plot 3.15b, the PWC values move away from the target as the damping rises. This would explain some of the changes in the error seen in Plot 3.15a, but raises the question as to whether the target level should be adjusted to compensate for changes to the damping coefficient. However, if the systems were equally effective at creating a plane wave, and the only issue was related to the gain from a non-ideal target, it would be expected that Plot 3.15c would show a flatter line, as even if the levels changed at each receiver point, the variance should have stayed the same. The MSV values in Plot 3.15c show that only when $\zeta_{\mathbf{n}} = \frac{27.631}{\omega_n}$ is the seat-to-seat variation in the unacceptable range but the plane wave cancellation method is always below the 15 dB threshold denoting a good level of seat-to-seat variation. As the damping rises, the system becomes less effective at creating a plane wave (at either front or rear), as the reinforcement to the sources created by the image method, as explained in Section 3.1, is not as strong. Therefore, the sources become more akin to individual spherical sources, leading to a degradation in the wave.



(a) Change in $\bar{\varepsilon}_{(10-100\,\text{Hz})}$.



(b) Change in RMS.



(c) Change in MSV.

FIGURE 3.15: Change in the $\bar{\varepsilon}_{(10-100\,\text{Hz})}$, RMS and MSV for different damping coefficients; RS (-●-RS), PWCs (-●-PWC) and target (-●-Target).

3.3.2 The effect on the plane wave cancellation method of moving The values of s_1 and s_2 perpendicular to the front wall

In this study, the sources s_1 and s_2 were moved as a pair from the ideal position on the front wall of the room in 0.1 m increments along the x-axis, from 0 m to 2 m, to produce 21 sets of data. The rear source and receiver locations did not change, and $\zeta_{\mathbf{n}}$ was set as $\frac{11.052}{\omega_n}$. In each test iteration, the level and phase of the signal from sources s_1 and s_2 were adjusted in order to compensate for the changes in distance and level created between the sources and the rear wall by moving the source away from the front wall. This was done in order to achieve the best possible cancellation at the rear surface. The change in distance could easily be compensated for by a change in phase, however the change in level was not straightforward, and it was chosen that the level would be adjusted to keep the same RMS pressure level at the rear boundary.

Figure 3.16 shows that when sources s_1 and s_2 were moved 0.7 m or more from the wall, the plane wave cancellation method lines started to cross the RS line in the worst

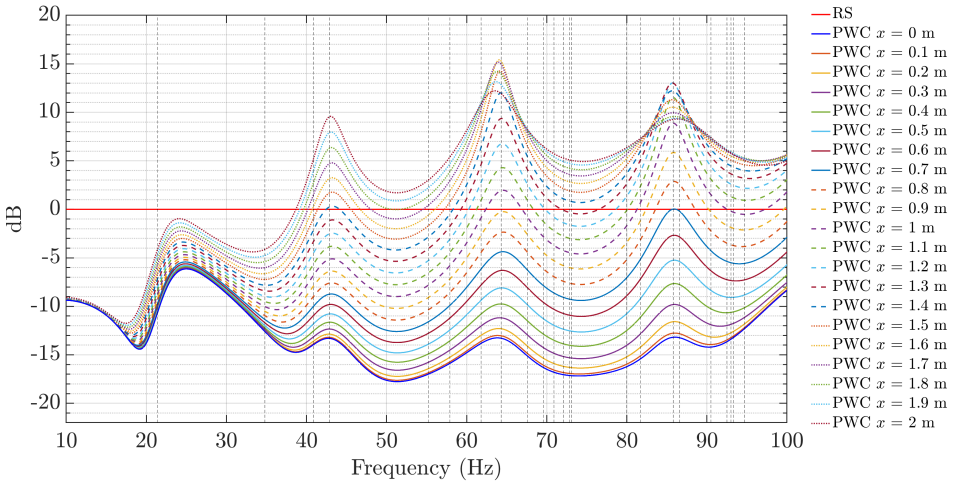
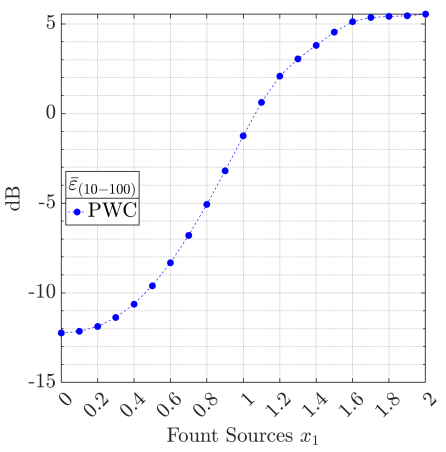


FIGURE 3.16: Error ratio against frequency for the RS (—RS) with no cancellation and 21 PWCs with different s_1 and s_2 locations measured over a fixed listening area. Ideal locations shown as (—), with all modal frequencies of the enclosure denoted as - -.

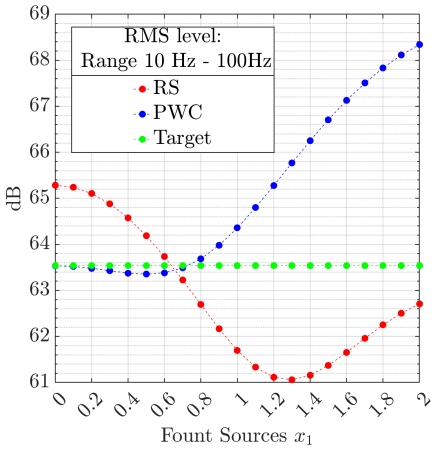
case, this starts as low as 39 Hz, and as the sources moved further away, the RS line was crossed more frequently. This meant that the plane wave cancellation method was less effective compared to doing nothing at this point. In Plot 3.17a, $\bar{\varepsilon}_{(10-100\,\mathrm{Hz})}$ averaged over the frequency clearly show that the effectiveness went down as the sources were moved away from the front wall; however, $\bar{\varepsilon}_{(10-100\,\mathrm{Hz})}$ when moved up to 0.4 m away, showed only a 1.5 dB drop in effectiveness, which increased to a 6 dB drop when moved up to 0.6 m away.

Plot 3.17b shows the change in the RMS level averaged over the frequency range as the source location changed. Unlike the previous study, the average RMS level of the raw signal is not the same for each test case. This is because the level was set to 63.53 dB over the listening area when the sources were on the wall. As the sources were moved away from the wall, the output level was dropped by an amount equivalent to the distance, so that the level over the listening area would stay the same. The changes that are seen in the average RMS level of the raw signal in Plot 3.17b are due to the change of source location affecting the relevant levels of the modes over the area. The MSV level of the RS was above 35 dB for all locations, and for the PWC methods the MSV dropped from good to acceptable at 0.8 m.

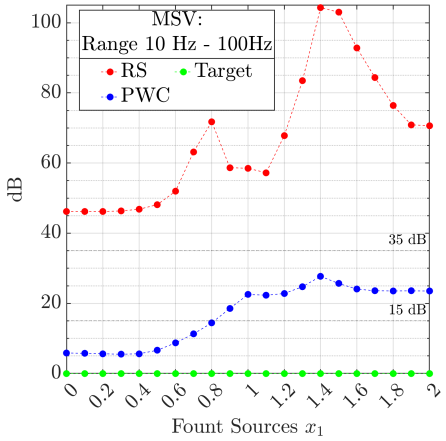
As expected, the average RMS level matched the target when the sources were in the ideal locations. Here, up to 0.8 m from the wall, there was a 1 dB difference between the target and the PWC. However, they started to diverge by up to 3 dB when they were moved up to 1.2 m. Again, this would explain some of the errors seen in Plot 3.17a. If the systems were all equally effective at creating a plane wave and the only issue was in the gain, it would be expected that Plot 3.17c would show a flatter line. Even if the levels changed at each receiver point, the variance would have stayed the same. The



(a) Change in $\bar{\varepsilon}_{(10-100\,\text{Hz})}$.



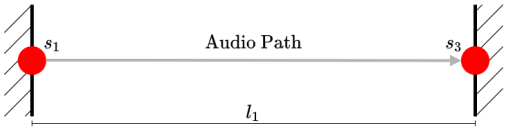
(b) Change in RMS.



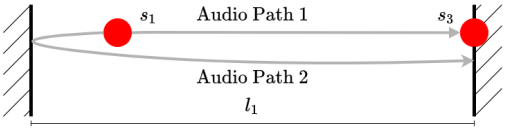
(c) Change in MSV.

FIGURE 3.17: Change in the $\bar{\varepsilon}_{(10-100\,\text{Hz})}$, RMS and MSV as the front sources are moved away from the front wall; RS (-●-RS), PWCs (-●-PWC) and target (-●-Target).

main error was created by the reflection from the front wall when the sources were not attached to it, as shown in Figure 3.18.



(a) Audio path sources are attached to the front wall.



(b) Audio path when sources are not attached to the front wall.

FIGURE 3.18: A representation of how moving the front source creates a second audio pathway and no cancelled reflections.

Figure 3.18 shows a representation of the tested system as viewed from the side; Plot 3.18a shows the ideal setup, and Plot 3.18b shows the source when moved off the front wall. Regarding the sources on the wall in Plot 3.18a, there was one audio path, which implied one wavefront would be cancelled at the rear wall. In Plot 3.18b, a second path

was created by moving the source away from the wall. In a standard plane wave cancellation method, only one of these audio paths can be cancelled (normally the first), which means that the second audio path signal will act as if there is no cancellation; this will be part of the error.

3.3.3 The effect on the plane wave cancellation method of moving s_1 and s_2 parallel to the front wall

In this study, the sources s_1 and s_2 were fixed on the front wall but moved along the y-axis. Each source was moved independently up to $\pm$ 1.0 m (in 0.1 m increments) from its ideal position along the same horizontal plane, giving 421 permutations for the pair of front sources. Again, the rear source and receiver locations did not change, and $\zeta_{\mathbf{n}} = \frac{11.052}{\omega_n}$.

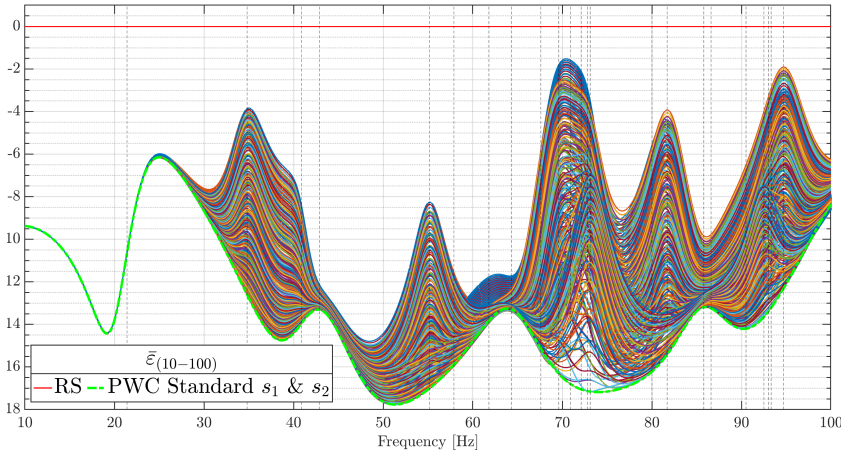


FIGURE 3.19: Error ratio against frequency for the RS (—RS) with no cancellation and 421 PWCs with different s_1 and s_2 locations measured over a fixed listening area. Ideal locations is shown as (- -), with all modal frequencies of the enclosure denoted as - -.

Figure 3.19 shows how the error ratio changed in the 421 sets of data. It is clear that there is too much data to analyse from this plot, but two observations can be made from Figure 3.19.

1. The PWC with ideal locations is shown as (- -), which was not the lowest of all frequencies but did track the minimal values with only a 2 dB difference at worst.

2. At no point were any of the plane wave cancellation method source permutations less effective than the RS.

Figure 3.20 shows the error $\bar{\varepsilon}_{(10-100\,\text{Hz})}$ averaged over the frequency as a surface plot, with the x-axis and y-axis showing the location of the sources s_1 and s_2. The z-axis shows the dB value for each pair of locations.

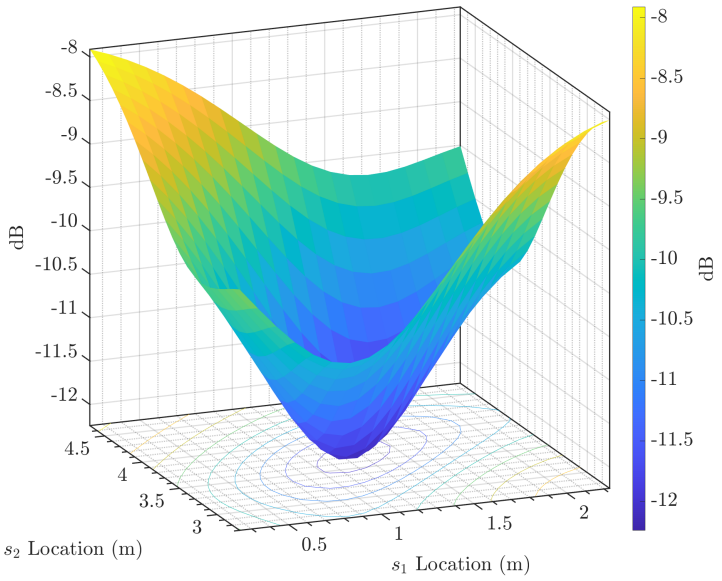


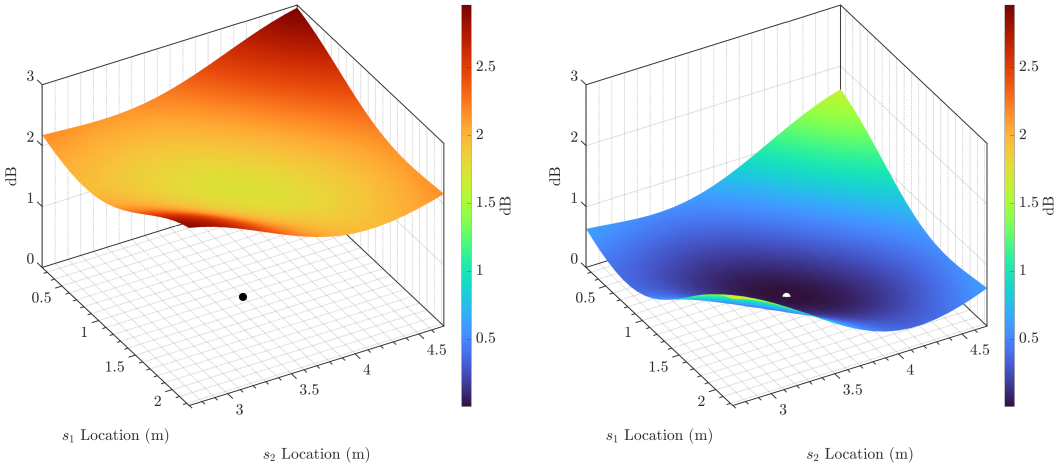
FIGURE 3.20: $\bar{\varepsilon}_{(10-100\,\text{Hz})}$ for 421 pairs of the s_1 and s_2 locations.

As shown in Figure 3.20, the minimum $\bar{\varepsilon}_{(10-100\,\text{Hz})}$ value was -12.24 dB; this was at the ideal source location. However, the maximum $\bar{\varepsilon}_{(10-100\,\text{Hz})}$ was -7.92 dB, and comparing this to the data in Section 3.3.2 shows that moving the source along the wall produced a smaller degradation in the system's effectiveness than moving the source away from the wall.

Figure 3.21 shows the $|\text{RMS}_\Delta|$ from the target for the RS (Plot 3.21a) and the $|\text{RMS}_\Delta|$ from the target for the PWC (Plot 3.21b). Figure 3.22 shows the MSV for the RS (Plot 3.22a) and the MSV for the PWC (Plot 3.21b). All with 421 permutations for the pair of front sources.

In this section, the RS values will be briefly discussed, with the focus on the change in effectiveness of the plane wave cancellation method. All plots in Figures 3.21 and 3.22 are surface plots, with the x-axis and y-axis giving the location of the source and the colour denoting the dB values at each pair of locations. In each case, the minimum value is shown by a dot, and in all but Plot 3.22a these minimum values are at the ideal source locations.

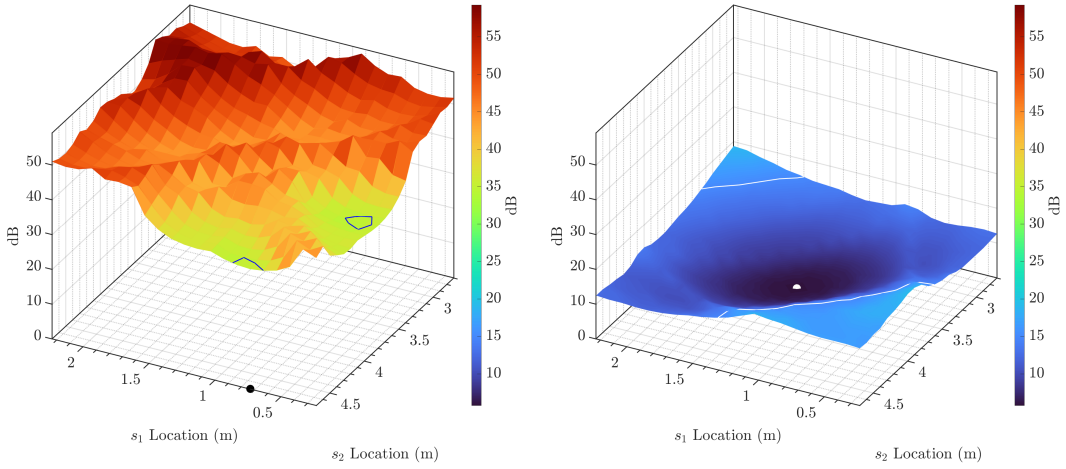
Figures 3.21 and 3.22 both show that the PWC methods offer an improvement towards the target over the RS, there is no location in either set of plots where the RS outperforms the PWC. In Plot 3.22a 98.87 % of the locations gave a MSV greater or equal to the



(a) RS $|\text{RMS}_\Delta|$. (b) RS $|\text{RMS}_\Delta|$.

FIGURE 3.21: $|\text{RMS}_\Delta|$ for RS and PWC with 421 pairs of the s_1 and s_2 locations.

35 dB threshold for acceptable seat-to-seat variation with a maximum MSV of 59.25 dB. There was a minimum MSV of 34.5 dB at the locations denoted by the black dot, and the only locations below 35 dB are indicated by the black contour lines, these locations are interesting but are irrelevant to the scope of this work.



(a) RS MSV. (b) PWC MSV.

FIGURE 3.22: MSV for RS and PWC with 421 pairs of the s_1 and s_2 locations.

In Plot 3.21a, the lowest value of $|\text{RMS}_\Delta|$ from the target is indicated by the white dot (this was at the ideal positions of s_1 and s_2) and was 0 dB. As the values moved away from this point, they increased to a maximum of 1.75 dB. Each colour step holds all values in a 0.25 dB banding. The $|\text{RMS}_\Delta|$ of the plane wave correction system was within 0.5 dB of the ideal positions for over 57 % of the positions and within 1 dB for 84.8 % of the positions.

In Plot 3.22b the closest MSV value to zero is indicated by the white dot (again at the ideal positions of s_1 and s_2) at 5.75 dB. All locations are below the 35 dB threshold, and the white contour lines indicate the 15 dB threshold between good to acceptable levels of seat-to-seat variation. 80.5% of the locations give a MSV below 15 dB, with 48.30 % of the positions within 6 dB of the minimum value and at a maximum MSV of 19.25 dB.

The data in Plots 3.21a and 3.22b again shows that moving the source along the wall produced less degradation on both the RMS and MSV values of the system compared to moving the source away from the wall. This was likely due to the issue created by moving the source on the front wall, as described by Figure 3.18.

3.3.4 Study review

The main outcomes of these studies were as follows:

- As the damping coefficient of the cuboid enclosure increased, the plane wave cancellation method was less effective.
- As the damping coefficient of the cuboid enclosure increased, the MSV values show that the RS could outperform the plane wave cancellation method.
- If the front sources were moved away from the front wall, the plane wave cancellation method was less effective.
- If the front sources were kept on the front wall but moved from their ideal locations, the plane wave cancellation method was less effective.
- The effectiveness of the plane wave cancellation method was more sensitive to the front source being moved away from the front wall than to the front source being moved along the front wall.

It has been shown in this section that the effectiveness of the plane wave cancellation method is reduced when the system setup is not ideal. This raises an opportunity for a system of control that is less restrictive in setup.

3.4 Chapter review

Having first explained the theory behind the plane wave cancellation method, this chapter has shown that although this method can offer high levels of effective control over the acoustic environment, there are limitations based on the source location and room characteristics. These limitations are mainly by-products of the theory discussed

in Section 3.1 and linked directly to what is both the simplicity and beauty of the plane wave cancellation method − namely, exploiting the fundamental physical properties of a cuboid environment to allow the principles of the image method and WFS to be utilised to create an effective solution that does not overly rely on large amounts of signal processing and offers control over the whole acoustic environment. However, the limitations that the location of the source and damping impose on the plane wave cancellation method systems make it hard to see how this approach could be robust enough to be effective in the real world. There are too many unforeseen environmental factors and little to no control over how the system will be set up and used.

Chapter 4

Application of the least-squares method for the control of room acoustics

The previous chapter showed that although the plane wave cancellation method can offer an effective level of control for the sound field at low frequencies within a cuboid room, this effectiveness degrades if the ridge setup requirements of the system are not followed, and the end user may struggle to adhere to such constraints in the home environment.

In this chapter, four different approaches for controlling the sound field over a fixed listening area within a cuboid environment are introduced, all using the least-squares method. These are then compared with the plane wave cancellation method. The four different processes aim to achieve acoustic control by simultaneously filtering up to four source signals with filters designed using the least-squares method. The theory is that by using the least-squares approach and adding a constraint to the listening area, it is possible to achieve an approach that can give as good or better results than the plane wave cancellation method but with more relaxed source locations and environmental constraints.

This chapter is split into the following sections:

1. introduction of the method of least-squares as a means of controlling room acoustics
2. comparison of the least-squares approaches with the plane wave cancellation method
3. real-world experimental study

4. multiple room layout testing
5. limitations of the least-squares approach

If the least-squares approach can create a plane wave sound field across the fixed listening area, the theory suggests that it may offer the possibility of redefining the target sound field. In order to gain a better understanding of this theory, a spherical wave target sound field is also investigated in this chapter.

4.1 Least-squares control methods

This section introduces the method of least-squares by providing a formulation of a solution in a general form as it pertains to this thesis. Afterwards, the specific approaches used in this thesis are explained, and each process of using the least-squares method for a filter design is presented.

4.1.1 Introduction of the method of least-squares

The method of least-squares is a widely established statistical regression tool to find the best fit for a set of data points by minimising the sum of the offsets or residuals of points from the target. It is used in audio, as in the *Active Control of Sound* by Nelson and Elliott (1992) and in Elliott and Nelson (1989), acoustic pressure matching (Kirkeby and Nelson (1993)), personal audio (Simón Gálvez et al. (2012)) and automotive audio (Cheer (2012)). In this thesis, the method of least-squares was used to create multiple-point, multiple-source digital filters for the control of room acoustics, which was then compared with the plane wave cancellation method. A general acoustic solution was expressed before moving on to specific cases.

Taking the source and receiver locations from Chapter 3 as $\mathbf{s}_w$ and $\mathbf{r}_m$, respectively, Figure 4.1 gives a representation of the system as a block diagram where the input signal is given as $v(f)$, which, unless otherwise stated, equates to the delta function in the time domain, and each source output signal is $\mathbf{u}_w(f)$.

Each element number in $\mathbf{u}_w(f)$ corresponds to the same element number $\mathbf{s}_w$, where $\mathbf{u}(f)$ is defined as

$$\mathbf{u}(f) = \mathbf{h}(f)v(f), \tag{4.1}$$

where $\mathbf{h}(f)$ are the calculated filters. In Figure 4.1, the transfer functions $\mathbf{s}_w \rightarrow \mathbf{r}_m$ are defined as the matrix $\mathbf{C}(f)$ of size $(M \times W)$, where M is the total number of receivers, and W is the total number of sources. At this stage, it is assumed the sources and

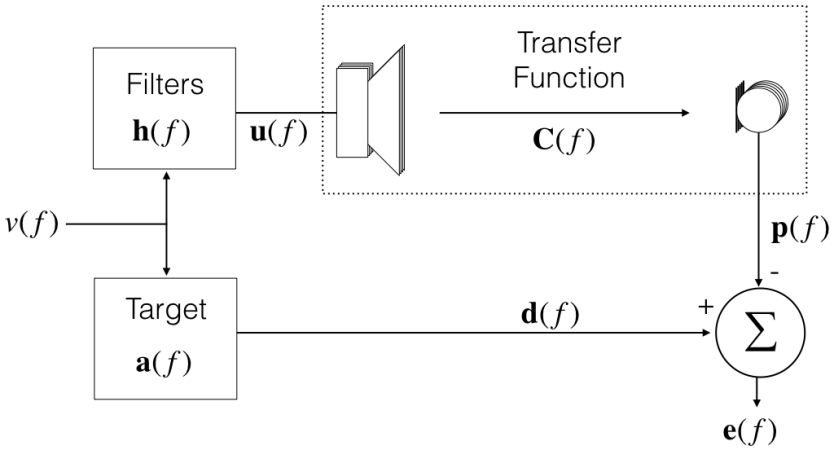


FIGURE 4.1: System block diagram.

receivers have no effect on the audio. The receivers' signal vector $\mathbf{p}(f)$ can be expressed as

$$\mathbf{p}(f) = \mathbf{C}(f)\mathbf{u}(f). \tag{4.2}$$

In Figure 4.1, $\mathbf{a}(f)$ defines the target transfer functions for each receiver point so that the pressure target can be defined as

$$\mathbf{d}(f) = \mathbf{a}(f)v(f). \tag{4.3}$$

Unless otherwise stated, it should be assumed that all variables are in the frequency domain, and therefore, in order to simplify the equations, the (f) notation is not used.

The aim is to calculate $\mathbf{h}$ such that

$$\mathbf{a}v = \mathbf{C}\mathbf{h}v. \tag{4.4}$$

Assuming that $v = 1$, this becomes

$$\mathbf{a} = \mathbf{C}\mathbf{h}. \tag{4.5}$$

The error can be defined as

$$\mathbf{e} = \mathbf{d} - \mathbf{p} \tag{4.6}$$

or

$$\mathbf{e} = \mathbf{a} - \mathbf{C}\mathbf{h}. \tag{4.7}$$

As already discussed, **a** represents the targets, **C** is the transfer function, and **h** is the filters that are calculated to minimise the error. The aim is to minimise the sum of the squared errors. To do this, the error criterion or cost function is, in general, defined as the sum of the square of the errors:

$$J = \sum_{m=1}^{M} |e_m|^2 = \mathbf{e}^{\mathrm{H}}\mathbf{e}. \tag{4.8}$$

By combining Equation 4.6 and Equation 4.8, a Hermitian quadratic equation is derived:

$$J = \mathbf{a}^{\mathrm{H}}\mathbf{a} - \mathbf{a}^{\mathrm{H}}\mathbf{C}\mathbf{h} - \mathbf{h}^{\mathrm{H}}\mathbf{C}^{\mathrm{H}}\mathbf{a} + \mathbf{h}^{\mathrm{H}}\mathbf{C}^{\mathrm{H}}\mathbf{C}\mathbf{h}. \tag{4.9}$$

As with a normal quadratic equation, this is solvable to find the minimum value. There are three cases:

Case 1 : $W > M$ The system is underdetermined, and the minimum-norm solution is used.

Case 2 : $W = M$ The system is fully determined.

Case 3 : $W < M$ The system is overdetermined, and the least-squares solution is used.

All cases in this thesis fall under **Case 3**, and the least-squares solution is used to minimise the error, which, as a solution to Equation 4.9, gives

$$\mathbf{h} = \left[\mathbf{C}^{\mathrm{H}}\mathbf{C}\right]^{-1}\mathbf{C}^{\mathrm{H}}\mathbf{a}. \tag{4.10}$$

4.1.2 Least-squares approaches for acoustic control

The least-squares method was used to create the filters for the four acoustic control approaches under investigation in this work. A graphical representation of the approaches is given in Figure 4.2. All four filter design methods were based on the plane wave cancellation method system setup. Up to four sources were available, and each approach required multiple point measurements over the fixed area to create the least-squares filters. The difference between each approach was the number of sources filtered at any one time. The four proposed approaches were

- **two-channel (front) approach** (room configuration Figure 4.2a): Only the front two sources controlled the sound field in the listening area. This was a multiple-point equalisation system, as described by Elliott and Nelson (1989).

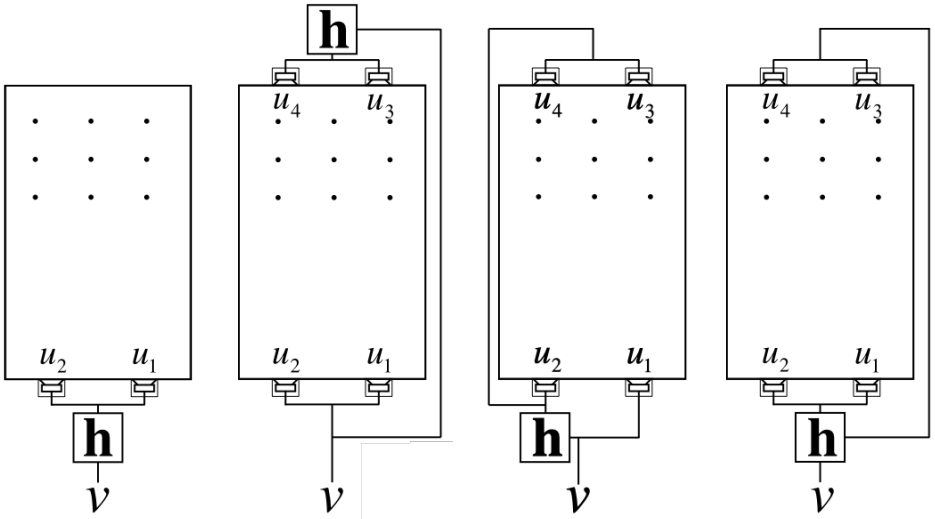


FIGURE 4.2: Room layout representation for each control approach [left to right: two-channel (front) approach, two-channel (rear) approach, three-channel approach and four-channel approach], with receiver grid represented as •.

- **two-channel (rear) approach** (room configuration Figure 4.2b): All four sources were used, but only the rear channels controlled the sound field in the listening area. This was akin to the plane wave cancellation method but with more freedom to find the best option to control the sound field rather than just adjusting the delay.
- **three-channel approach** (room configuration Figure 4.2c): All four sources were used, but only three channels controlled the sound field in the listening area, with one primary source left unfiltered. When evaluating the spherical wave target sound field, the spherical wave was propagated from the primary source.
- **four-channel approach** (room configuration Figure 4.2d): This solution used all four sources to control the sound field in the listening area.

All the results were compared with the plane wave cancellation method over a fixed listening area, as introduced in Chapter 3, and all of this work was conducted at frequencies below 100 Hz.

4.1.2.1 Least-squares filter design

As stated in Section 4.1.1, all the filters were overdetermined, and a least-squares solution was used. However, each approach was different in the way it applied the solution.

The four-channel and two-channel (front) approaches applied the least-squares solution directly, as given in Equation 4.10, and the difference was the number of channels. The four-channel system used all four sources (s_1, s_2, s_3 and s_4), and therefore, the filters were in the form of

$$\mathbf{h} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix}. \tag{4.11}$$

By contrast, the two-channel (front) system used sources s_1 and s_2 only, which meant the filter was in the form

$$\mathbf{h} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}. \tag{4.12}$$

The difference was in how Equation 4.2 was defined. With the four-channel system, Equation 4.2 was defined as

$$\mathbf{p} = \mathbf{C}\mathbf{u} = \begin{bmatrix} \hat{\mathrm{P}}(r_1|s_1) & \hat{\mathrm{P}}(r_1|s_2) & \cdots & \hat{\mathrm{P}}(r_1|s_4) \\ \hat{\mathrm{P}}(r_2|s_1) & \hat{\mathrm{P}}(r_2|s_2) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathrm{P}}(r_M|s_1) & \cdots & \cdots & \hat{\mathrm{P}}(r_M|s_4) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_4 \end{bmatrix}. \tag{4.13}$$

However, in the two-channel (front) system, Equation 4.2 is defined as

$$\mathbf{p} = \mathbf{C}\mathbf{u} = \begin{bmatrix} \hat{\mathrm{P}}(r_1|s_1) & \hat{\mathrm{P}}(r_1|s_2) \\ \hat{\mathrm{P}}(r_2|s_1) & \hat{\mathrm{P}}(r_2|s_2) \\ \vdots & \vdots \\ \hat{\mathrm{P}}(r_M|s_1) & \hat{\mathrm{P}}(r_M|s_2) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{4.14}$$

Both could still be expressed in the form

$$\mathbf{a} = \mathbf{C}\mathbf{h}, \tag{4.15}$$

which led to

$$\mathbf{h} = \left[\mathbf{C}^{\mathrm{H}}\mathbf{C}\right]^{-1}\mathbf{C}^{\mathrm{H}}\mathbf{a}, \tag{4.16}$$

as laid out in Equation 4.10.

The two-channel (rear) solution was a four-source system that used s_1, s_2, s_3 and s_4 but only filters s_3 and s_4. Therefore, the filter was in the form of

$$\mathbf{h} = \begin{bmatrix} 1 \\ 1 \\ h_3 \\ h_4 \end{bmatrix}. \tag{4.17}$$

The pressures vector (Equation 4.2) could be rewritten as the sum of the front ($\mathbf{p}_f$) and rear ($\mathbf{p}_b$) pressures,

$$\mathbf{p} = \mathbf{p}_f + \mathbf{p}_b, \tag{4.18}$$

where

$$\mathbf{p}_f = \begin{bmatrix} \hat{\mathrm{P}}(r_1|s_1) & \hat{\mathrm{P}}(r_1|s_2) \\ \hat{\mathrm{P}}(r_2|s_1) & \hat{\mathrm{P}}(r_2|s_2) \\ \vdots & \vdots \\ \hat{\mathrm{P}}(r_M|s_1) & \hat{\mathrm{P}}(r_M|s_2) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \qquad \mathbf{p}_b = \begin{bmatrix} \hat{\mathrm{P}}(r_1|s_3) & \hat{\mathrm{P}}(r_1|s_4) \\ \hat{\mathrm{P}}(r_2|s_3) & \hat{\mathrm{P}}(r_2|s_4) \\ \vdots & \vdots \\ \hat{\mathrm{P}}(r_M|s_3) & \hat{\mathrm{P}}(r_M|s_4) \end{bmatrix} \begin{bmatrix} u_3 \\ u_4 \end{bmatrix} \tag{4.19}$$

or

$$\mathbf{p}_f = \mathbf{C}_f \mathbf{u}_f \qquad \mathbf{p}_b = \mathbf{C}_b \mathbf{u}_b, \tag{4.20}$$

where $\mathbf{u}_f$ was fixed as

$$\mathbf{u}_f = \begin{bmatrix} 1 \\ 1 \end{bmatrix} v. \tag{4.21}$$

Therefore,

$$\mathbf{p} = \mathbf{p}_f + \mathbf{C}_b \mathbf{u}_b \tag{4.22}$$

or

$$\mathbf{a} = \mathbf{p}_f + \mathbf{C}_b \mathbf{h}_b, \tag{4.23}$$

leading to

$$\mathbf{h}_b = \mathbf{C}_b^{\dagger}(\mathbf{a} - \mathbf{p}_f). \tag{4.24}$$

Therefore,

$$\mathbf{h}_b = \left[\mathbf{C}_b^{\mathrm{H}} \mathbf{C}_b\right]^{-1} \mathbf{C}_b^{\mathrm{H}}(\mathbf{a} - \mathbf{p}_f) \tag{4.25}$$

and

$$\mathbf{h}_f = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{4.26}$$

$$\mathbf{h} = \begin{bmatrix} \mathbf{h}_f \\ \mathbf{h}_b \end{bmatrix}. \tag{4.27}$$

As with the two-channel (rear) solution, the three-channel control method was a four-source system using s_1, s_2, s_3 and s_4 with one primary source left unfiltered, and three secondary sources controlled the sound field. Therefore, the filter was in the form

$$\mathbf{h} = \begin{bmatrix} 1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix}. \tag{4.28}$$

This time, Equation 4.2 was rewritten as the sum of s_1 ($\mathbf{p}_\alpha$) and s_2, s_3 and s_4 ($\mathbf{p}_\beta$) pressures:

$$\mathbf{p} = \mathbf{p}_\alpha + \mathbf{p}_\beta. \tag{4.29}$$

By following similar steps as outlined in the two-channel (rear) solution, the three-channel control least-squares solution was given by

$$\mathbf{h}_\beta = \left[\mathbf{C}_\beta^{\mathrm{H}}\mathbf{C}_\beta\right]^{-1}\mathbf{C}_\beta^{\mathrm{H}}(\mathbf{a} - \mathbf{p}_\alpha) \tag{4.30}$$

and

$$\mathbf{h}_\alpha = 1 \tag{4.31}$$

with the filters given as

$$\mathbf{h} = \begin{bmatrix} \mathbf{h}_\alpha \\ \mathbf{h}_\beta \end{bmatrix}. \tag{4.32}$$

In this section, the specific solutions for the design of each set of filters for the four least-squares control approaches have been given. These will now be used throughout the thesis.

4.1.2.2 Least-squares filter implementation

The previous section outlined how the least-squares filters have been designed for each approach, in each case giving a vector $\mathbf{h}(f)$ in the frequency domain. These vectors hold the magnitude and phase specification to best achieve the defined target in the given environment, and each vector is made up of equally spaced samples along the frequency axis. Throughout this thesis, in both the numerical model and real-world testing, the filters are not applied to the test signal in real-time; the test signal is processed offline, meaning there is less restriction on processing power, allowing the resolution of these samples to be 10^{-2} Hz in the numerical modal and 0.25 Hz in real-world testing. The most straightforward way to create audio from this vector is to use the frequency sampling approach (Oppenheim et al. (1982)), which is used to design a finite impulse response (FIR) filter in the time domain. The combined magnitude and phase

specifications from a complex frequency response representation give the filters' impulse response by applying the inverse discrete Fourier transform. This results in the desired impulse response of the FIR filter Smith (2003). The filter can also be applied in the frequency domain by taking the discrete Fourier transform of the signal to be filtered, multiplying the two frequency vectors together, and then performing the inverse discrete Fourier transform. One issue with this method is that the filter is undefined between each sample point. The frequency sampling approach is common in sound field control, and even though it is not always implicitly stated, it can be seen in use in many texts, including Olivieri et al. (2013), Kirkeby et al. (1996), Nelson and Elliott (1988) and Kirkeby and Nelson (1999).

4.1.2.3 Hypotheses

The expectations are

- **two-channel (front) approach**: This will be the worst-performing approach. This type of method had been investigated in the past by Elliott and Nelson (1989). Here, it is being used as a second benchmark. If the approaches with more sources cannot outperform a two-source method, there is no point in the effort needed with more sources.
- **two-channel (rear) approach**: This approach will always be as good as or better than the plane wave cancellation method, at least in the controlled region, because the least-squares approach will find the best possible solution. Therefore, if the plane wave cancellation method is the best possible solution, the two-channel (rear) control will converge on it.
- **three-channel approach**: This approach will always be as good as or better than the two-channel (rear) approach for the same reasons that two-channel (rear) control cannot be worse than the plane wave cancellation method. The extra degree of freedom means that, at worst, it will converge on the two-channel (rear) solution.
- **four-channel approach**: This approach will always be as good as or better than any other method discussed in this work. It has the highest number of degrees of freedom to converge on the best possible solutions. Therefore, if one of the other methods under investigation finds the best possible solution, this method will also find that solution.

4.2 Comparison of the least-squares approaches with the plane wave cancellation method

This section assesses the effectiveness of the four least-squares control approaches against the plane wave cancellation method reviewed in Chapter 3, through two studies with a fixed listening area. Each study used the analytical model from Chapter 3 in which the transfer functions were created using the Green's function previously given in Equation 2.23 as

$$\hat{\mathrm{P}}(\mathbf{r}|\mathbf{s}) = \sum_{n=0}^{\infty} \frac{\omega\rho_0 c_0^2 Q\ \psi_n(\mathbf{r})\psi_n(\mathbf{s})}{[2\zeta_n\omega_n\omega + i(\omega^2 - \omega_{\mathbf{n}}^2)]\mathrm{V}} \tag{4.33}$$

with $\zeta_{\mathbf{n}} = \frac{11.052}{\omega_n}$. As with Section 3.2.2, the parameters of the investigations were set to obtain the most effective outcome from the plane wave cancellation method. Again, the model was based on the Bowers & Wilkins SRE main listening room with dimensions of $8.00 \times 4.93 \times 2.96$ m. The source and receiver locations were the same as those given in Section 3.2 in Tables 3.1 and 3.6, respectively.

4.2.1 Method comparison with the plane wave target sound field

This study repeated the work carried out in Section 3.2.4, with the addition of the least-squares methods. The target sound field over the listening area was a plane wave, previously given as

$$p(x) = p_0 e^{-ik\mathbf{r}_m}. \tag{4.34}$$

The plane wave target was travelling through the listening area, again propagating parallel to the front surface, where the level had been set the same as in Section 3.2.4 to give the best performance from the plane wave cancellation method.

4.2.1.1 Results

Figure 4.3 shows the error ratio ε as defined in (2.32) for all methods across the listening area for the frequency range of $10-100$ Hz, with a plane wave target sound field, where the raw signal (—RS) at 0 dB gave a normalised reference point on which the other methods aimed to improve. It was clear that all methods in Figure 4.3 showed a level of improvement over the raw signal. As predicted in Section 4.1.2.3, the four-channel least-squares approach (—4C) was the most effective, but was tracked closely by the three-channel least-squares approach (—3C). The two-channel (rear) least-squares approach (- - $2C_R$) was less effective than the three-channel approach but more effective

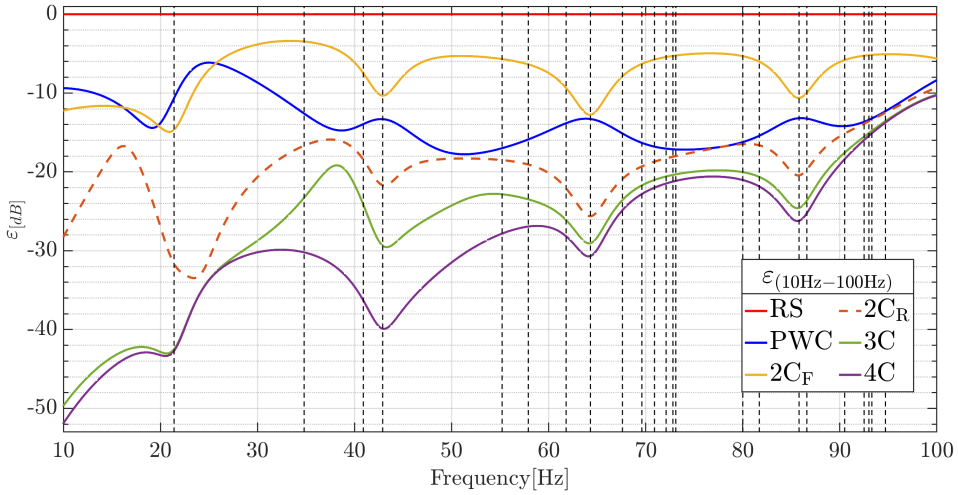


FIGURE 4.3: Error ratio ε against the frequency of the raw signal with no cancellation (—RS), of the plane wave cancellation method (—PWC), two-channel (front) approach (—2C$_F$), two-channel (rear) approach (- - 2C$_R$), three-channel approach (—3C) and four-channel approach (—4C), over the fixed listening area with a plane wave target, with all the modal frequencies of the enclosure denoted as - -.

than the plane wave cancellation method. The two-channel (rear) approach very closely tracked the three-channel approach above 55 Hz. The two-channel (front) approach was less effective than the plane wave cancellation method.

Table 4.1 shows the frequency band $\bar{\varepsilon}$, $|\text{RMS}_\Delta|$ from the target, and the MSV for all methods across the listening area with a plane wave target sound field as defined in Section 3.2.4. The error data ($\bar{\varepsilon}$) in Table 4.1 endorsed that shown in Figure 4.3 in the full frequency range (10 to 100 Hz). The four-channel, three-channel and two-channel (rear) approaches were all more effective than the plane wave cancellation method, showing improvements of 9.08, 8.09 and 5.21 dB, respectively. The two-channel (front) approach was 5.76.2 dB worse than the plane wave cancellation method. As can be seen in Figure 4.3, as the frequency increased, the delta between the methods dropped. In the upper frequency range (70 to 100 Hz), the four-channel, three-channel and two-channel (rear) approaches were still more effective than the plane wave cancellation method, with lower levels of improvement of 3.29, 2.91 and 1.21 dB, respectively.

In Table 4.1, the $|\text{RMS}_\Delta|$ from the target showed a 0 dB value for the plane wave cancellation method across the full range, as discussed in Chapter 3. This was due to the target RMS level being specifically chosen in order to be achievable by the plane wave cancellation method in this setup. The four-channel, three-channel and two-channel (rear) approaches were all, at worst, within 1.43 dB of the target, with the two-channel (front) showing the worst level of any correction method.

Table 4.1 also shows the band-limited data for MSV and this correlated with the data

TABLE 4.1: Frequency banded mean error [dB], RMS [dB] and MSV [dB] over the fixed listening area with a plane wave target.

Method	10 Hz $> f \geq$ 100 Hz			10 Hz $> f \geq$ 40 Hz			40 Hz $> f \geq$ 70 Hz			70 Hz $> f \geq$ 100 Hz		
	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV
RS	0.00	2.52	46.17	0.00	2.14	23.93	0.00	2.67	47.72	0.00	2.72	67.00
PWC	-12.24	0.00	5.85	-9.80	1.73	2.73	-15.13	1.21	7.85	-13.62	1.26	6.99
$2C_F$	-6.48	1.35	46.27	-6.51	2.29	23.81	-7.12	0.99	48.21	-5.88	0.57	66.96
$2C_R$	-17.45	0.93	1.78	-19.77	0.13	0.82	-19.74	1.36	0.84	-14.83	1.43	3.69
3C	-20.33	0.26	1.26	-26.67	0.10	0.17	-24.65	0.28	0.44	-16.53	0.40	3.17
4C	-21.32	0.08	1.03	-33.60	0.00	0.02	-28.69	0.02	0.25	-16.91	0.23	2.81

already seen in this section, attesting that the four-channel, three-channel and two-channel (rear) approaches were all more effective than the plane wave cancellation method. However, all these methods are below the 15 dB threshold that demonstrates good system performance. The two-channel (front) approach was less effective than the plane wave cancellation method with the MSV above the 35 dB threshold and therefore outside the acceptable range of system performance, showing little change from the raw signal.

Figure 4.4 shows the SA for all methods across the listening area for the frequency range of 10−100 Hz with a plane wave target sound field. Up to 80 Hz, all the least-squares approaches were closer to the target than the raw signal at the points where the raw signal had peaks and troughs. In many places, the four-channel, three-channel and two-channel (rear) approaches were so close to the target that they could not be distinguished from each other. All the data so far build a picture that the four-channel, three-channel and two-channel (rear) approaches were more effective than the plane wave cancellation method.

4.2.2 Method comparison with the spherical wave target sound field

This study moved away from defining the target sound field as a plane wave and instead defined the desired sound field over the listening area as a spherical wave given as

$$p(\mathbf{r}_m) = p_0 \frac{e^{-ik||\mathbf{s}_w - \mathbf{r}_m||}}{4\pi||\mathbf{s}_w - \mathbf{r}_m||}, \tag{4.35}$$

where, again, p_0 was set to the same level as in Section 3.2.4 wherein $||\mathbf{s}_w - \mathbf{r}_m||$ was the direct path in metres between the source and the receiver point combination under test. The target at each measurement location was a single spherical wave propagating from s_1. This target was akin to removing all wall reflections as if the sources were active in the free field.

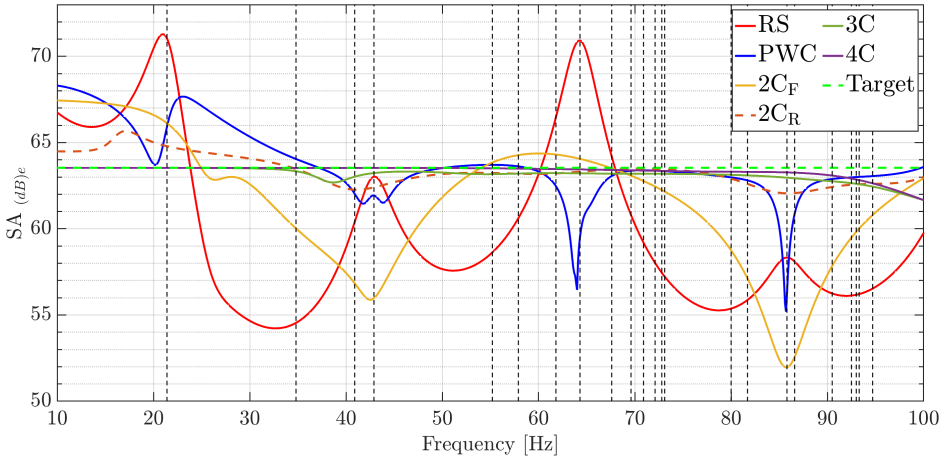


FIGURE 4.4: SA against the frequency of the raw signal with no cancellation (—RS), the plane wave cancellation method (—PWC), two-channel (front) approach (—$2C_F$), two-channel (rear) approach (- - $2C_R$), three-channel approach (—3C) and four-channel approach (—4C), over the fixed listening area with a plane wave target, with all the modal frequencies of the enclosure denoted as - -.

4.2.2.1 Why a spherical wave target?

The plane wave cancellation method is, by design, limited to creating a plane wave in a listening environment, so changing the target may seem unjustified. One reason to investigate another target is to understand the least-squares approaches' limitation. If the least-squares approaches are as effective at achieving a spherical wave target as a plane wave target, there can be a discussion (outside the remit of this thesis) on using these approaches to impose any type of target over a fixed listening area.

Another reason to investigate a spherical wave target is that it is possible that the two target types would have different focuses when transferring this acoustic control to a home listening environment:

- A plane wave target over the listening area is for a stereo listening experience when creating a soundstage in front of the user.
- A spherical wave target over the listening area is for a multi-channel setup (5.1, 7.1, etc.) wherein individual loudspeakers may have no shared content, and the audio can engulf the user from any direction.

4.2.2.2 Results

Figure 4.5 shows the error for all methods across the listening area for the frequency range of 10−100 Hz, and Table 4.2 shows the band-limited $\bar{\varepsilon}$, $|\mathrm{RMS}_\Delta|$ from the target,

and the MSV for all methods across the listening area with a spherical wave target sound field.

The first thing to record is that both Figure 4.5 and Table 4.1 show the effectiveness of the plane wave cancellation method (—PWC). Even though the plane wave cancellation method was designed to create a plane wave, the resultant sound field over the listening area seen in the data here, was closer to the spherical wave than the raw signal (—RS) and the two-channel (front) approach (—$2C_F$). One reason for this was the wavelength of the frequency under control. If the wavelength is wide enough compared with the total area of spatial sampling, the difference between a plane wave and a segment of a spherical wave propagated through that area is small.

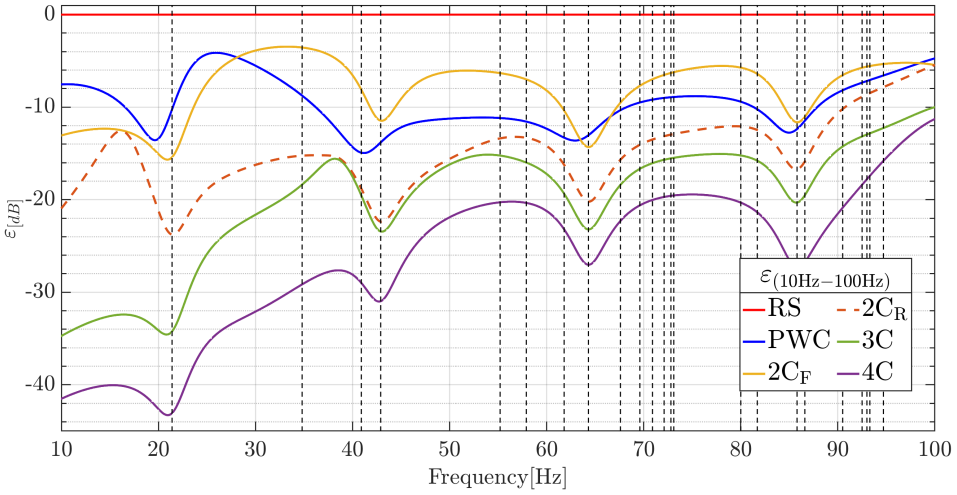


FIGURE 4.5: Error ratio ε against the frequency for the raw signal with no cancellation (—RS), of the plane wave cancellation method (—PWC), two-channel (front) approach (—$2C_F$), two-channel (rear) approach (- - $2C_R$), three-channel approach (—3C) and four-channel approach (—4C), over the fixed listening area with a spherical wave target, with all the modal frequencies of the enclosure denoted as - -.

Both the error data in Figure 4.5 and Table 4.2 are as seen in Section 4.2.1.1 and predicted in Section 4.1.2.3, namely that the four-channel least-squares approach (—4C) was the most effective, followed by the three-channel least-squares approach (—3C) and then the two-channel (rear) least-squares approach (- - $2C_R$). Although the plotted shapes of these three methods were similar, they did not track each other as closely as in Section 4.2.1.1, and the delta between the two-channel (rear) approach and the plane wave cancellation method, was smaller than in Section 4.2.1.1. Across the full frequency range, the four-channel approach was 4.37 dB more effective than the three-channel least-squares approach, 8.17 dB more effective than the two-channel (rear) approach and 12.48 dB more effective than the plane wave cancellation method.

The $|\text{RMS}_\Delta|$ from the target values in Table 4.2 show that for this target, the plane wave cancellation method was not the most effective over the full frequency band. Instead,

TABLE 4.2: Frequency banded mean error [dB], RMS [dB] and MSV [dB] over the fixed listening area with a spherical wave target.

Method	10 Hz $> f \geq$ 100 Hz			10 Hz $> f \geq$ 40 Hz			40 Hz $> f \geq$ 70 Hz			70 Hz $> f \geq$ 100 Hz		
	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV
RS	0.00	1.75	44.63	0.00	2.14	22.40	0.00	2.67	46.18	0.00	0.01	65.47
PWC	-8.59	0.00	4.32	-7.52	1.73	1.19	-11.22	1.21	6.32	-7.88	1.26	5.45
$2C_F$	-6.77	0.42	43.42	-7.03	0.20	24.23	-7.26	0.48	38.19	-6.10	1.10	67.99
$2C_R$	-12.90	1.73	1.94	-16.37	1.14	-0.24	-15.46	2.01	0.48	-9.89	2.11	5.59
3C	-16.74	2.38	1.26	-21.94	1.70	0.01	-17.56	2.55	0.47	-14.01	2.99	3.30
4C	-21.07	1.79	0.45	-33.52	1.70	-0.06	-22.80	1.79	-0.50	-17.50	1.89	1.92

it matched the other data in which the four-channel approach was the most effective, followed by the three-channel approach and then the two-channel (rear) approach. It was also one of the few data sets wherein the two-channel (front) approach offered comparable performance to the plane wave cancellation method. The MSV data in Table 4.2 correlated with the data already recorded in this section, wherein the four-channel, three-channel and two-channel (rear) approaches were more effective than the plane wave cancellation method, but these four methods are all below the 15 dB threshold value, in all data sets. Again, the two-channel (front) approach was less effective than the plane wave cancellation method with MSV values above the 35 dB threshold.

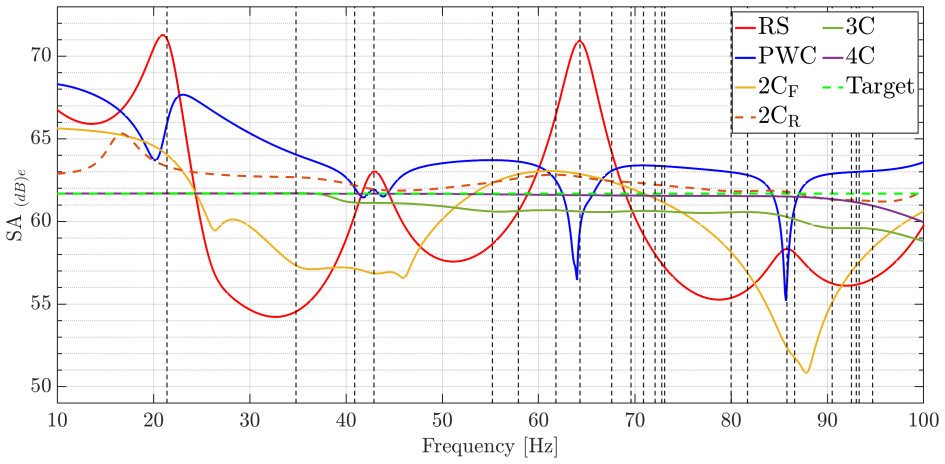


FIGURE 4.6: SA against the frequency of the raw signal with no cancellation (—RS), the plane wave cancellation method (—PWC), two-channel (front) approach (—$2C_F$), two-channel (rear) approach (- - $2C_R$), three-channel approach (—3C) and four-channel approach (—4C), over the fixed listening area with a spherical wave target, with all the modal frequencies of the enclosure denoted as - -.

Figure 4.6 shows the SA for all methods across the listening area for the frequency range of 10−100 Hz with a spherical wave target sound field. The target (- -) had a lower level than Figure 4.4. This was due to the nature of spherical wave propagation

and the averaging of the target level differences across the listening area. It is clear that only the four-channel and three-channel approaches coped with the gain change across the listening area, as they were indistinguishable from the target below 55 Hz.

Cross-comparing the data with a spherical wave target and a plane wave target showed an interesting result. Focusing on the $\bar{\varepsilon}$ data across the full band within Tables 4.2 and 4.2, all the approaches were less effective at achieving a spherical wave target than a plane wave target. The two-channel (rear) showed the biggest delta of 4.55 dB in effectiveness at achieving each target with the three-channel approach next biggest at 3.63 dB. The other methods showed much smaller deltas; 0.32 dB for the two-channel (front) approach and 0.25 dB for the four-channel approach. It is theorised that the reason for these differences was the room's set-up. As discussed in Chapter 3, the physical attributes of the environment were chosen specifically for plane wave cancellation, and two pairs of sources were placed to naturally create the best possible plane waves propagating from both back and front. When targeting a spherical wave over the listening area, the systems tried to work against these fixed physical properties. It is only logical that a system that has full control of all available sources (the two-channel (front) and four-channel approaches) will give good results. With the two-channel (rear) approach, the front sources naturally created a plane wave that the rear sources tried to modify into a spherical wave, explaining why this method was substantially less effective at achieving a spherical wave target.

4.2.3 Study summary

The results showed the hypotheses in Section 4.1.2.3 to be true, and from the studies in this section, the following statements can be made:

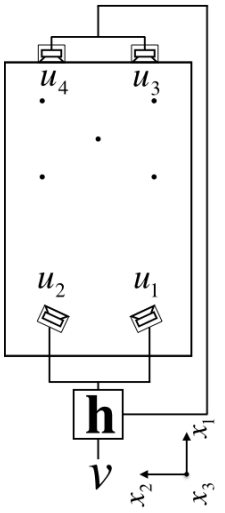
- The four-channel least-squares approach is the most effective, followed by the three-channel least-squares approach and then the two-channel (rear) least-squares approach.
- The two-channel (front) approach is not only the worst-performing of the least-squares approaches, but it is also worse than the plane wave cancellation method.
- When a room is set up to be ideal for the plane wave cancellation method, it is always more effective to use the four-channel, three-channel or two-channel (rear) least-squares approach than the plane wave cancellation method.
- When targeting a spherical wave sound field over the listening area, the plane wave cancellation method is more effective than the raw signal in order to achieve the required sound field. However, it is more effective to use the four-channel, three-channel or two-channel (rear) least-squares approach than the plane wave cancellation method.

As shown with the analytical model, the least-squares approach can improve the plane wave cancellation method. In the rest of this chapter, the least-squares approaches continue to be tested to discover how they transfer to the real world and are affected by the system layouts, to get a fuller picture of the control and possible limitations.

4.3 Real-world experimental studies

With the effectiveness of the least-squares approach proved with the analytical model, this section shows whether these results can transfer to the real world. If the results transfer, it will also verify the analytical model as a trusted platform on which to base further testing. Several measurements were carried out with different systems, targets and control approaches to determine what needed focus and what ideas should be carried forward or dropped. A summary of two studies is provided here, which compares real-world data with the analytical model for the four-channel least-squares approach. This section contains a short subsection outlining the changes made to the testing system needed when moving from a numerically modelled approach to real-world testing before giving the results for the last two studies.

All measurements were taken in the Bowers & Wilkins main listening room at the now closed SRE. This was the same room modelled in Section 4.2 with the dimensions of $8.00 \times 4.93 \times 2.96$ m. Unlike Section 4.2, the front sources were in a stereo-style layout, as it was decided to use a more real-world layout than the plane wave layout. The locations of the loudspeakers are given in Table 4.3.



(a)

FIGURE 4.7: Real-world testing room layout.

During the pre-testing for this work, it was decided that five receiver positions would be enough Bell and Fazi (2019). The locations of the microphones are given in Table 4.4.

TABLE 4.3: Real-world testing loudspeaker source location.

Source No.	Source Location ($\tilde{x}$ m, $\tilde{y}$ m, $\tilde{z}$ m)	Source No.	Source Location ($\tilde{x}$ m, $\tilde{y}$ m, $\tilde{z}$ m)
$\mathbf{s}_1$	(2.0, 1.3, 1.15)	$\mathbf{s}_3$	(7.9, 1.7, 1.15)
$\mathbf{s}_2$	(2.0, 3.63, 1.15)	$\mathbf{s}_4$	(7.9, 3.23, 1.15)

TABLE 4.4: Real-world testing microphone receiver location.

Receiver No.	Receiver Location (x m, y m, z m)	Receiver No.	Receiver Location (x m, y m, z m)
$\mathbf{r}_1$	(5.0, 1.65, 1.0)	$\mathbf{r}_4$	(7.0, 1.65, 1.0)
$\mathbf{r}_2$	(5.0, 3.65, 1.0)	$\mathbf{r}_5$	(7.0, 3.65, 1.0)
$\mathbf{r}_3$	(6.0, 2.65, 1.0)		

In this section, only the four-channel approach was tested with the created filter as defined in Section 4.1.2.1, as this had been proved to be the most effective approach. The studies followed these steps:

Step 1 − Measure the individual loudspeaker to microphone transfer functions. A swept sine-wave from 10 to 400 Hz was used to measure the signal from each loudspeaker to each measurement point. These measurements were then convolved with the inverse of the swept sine to create an impulse response.

Step 2 − Create the target. The required target was created in the band-limited region of 10−200 Hz.

Step 3 − Calculate the filter. The target and measured transfer functions were combined as outlined in Equation 4.10.

Step 4 − Calculate the estimated result. The filters were applied to the measured transfer functions in order to estimate the amount of improvement possible.

Step 5 − Filters were applied to the sine sweep, and new measurements were taken to obtain both real-world and modelled results.

4.3.1 System improvement

When moving from a numerically modelled approach to real-world testing, the system required some basic changes to improve its real-world performance; these improvements outlined in this section were integrated into the system.

4.3.1.1 Improve impulse calculation

In the simplest form, the impulses were created by convolving the capture signal at the measurement point with the inverse of the sweep. Figure 4.8 shows a single impulse that made up the full matrix of transfer functions used to create the correction filters.

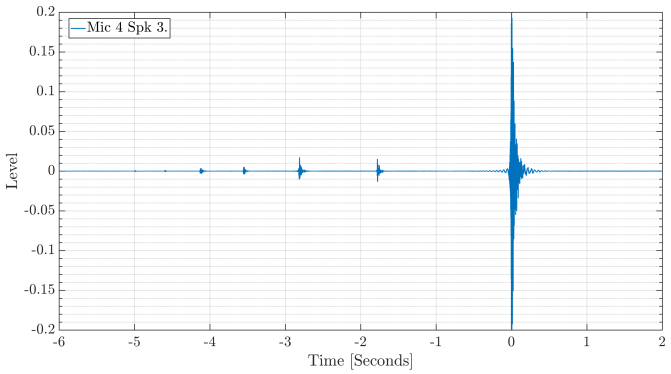


FIGURE 4.8: Single impulse response from loudspeaker three to microphone four.

In Figure 4.8, it is clear that there were numerous small peaks (pre-ringing) before the main peak of the impulse responses. The cause of this was harmonic distortion in the loudspeaker and the enclosure itself. In Farina (2000), these issues were quantified by the equation

$$\tau_n = T_n \frac{\ln(n)}{\ln(\frac{\omega_2}{\omega_1})}, \tag{4.36}$$

where T_n is the duration of the sweep used to make the measurement, ω_1 is the start frequency of the sweep, ω_2 is the stop frequency of the sweep, n is the order of the harmonic distortion, and τ_n is the time at which the peak due to the harmonic distortion appears before the main impulse. A longer sweep or lower stop frequency increases the distance between the main impulse and the harmonics.

To address the issue seen in Figure 4.8, a window was added in the time domain to remove the pre-ringing but keep the relative phase after testing. Choosing a one-second Tukey (tapered cosine) window with a 25 % cosine fraction gave the best result overall.

4.3.1.2 Filter target

The target design did not take into consideration any other limiting factors within the system. Figure 4.9 shows the frequency response of the loudspeakers used. These measurements were taken in the Bowers & Wilkins SRE anechoic chamber with an output of 2.84 volts at the terminals. The loudspeakers were quite evenly matched. All the measurements showed a drop in the sound pressure level under 40 Hz, meaning that before trying to correct any room issues in this range, filters had to correct the

loudspeaker first. Although the aim of the system was not to correct the loudspeaker directly, it was wise to take the loudspeaker into account when designing the filters.

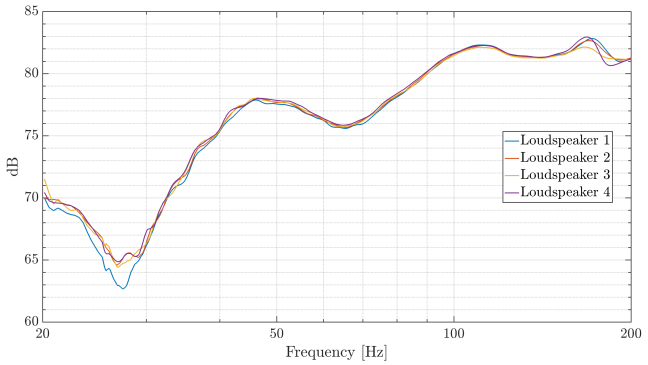


FIGURE 4.9: Loudspeaker frequency responses.

Figure 4.10 shows the frequency response of the bandpass filter applied to the target by taking the loudspeakers into account.

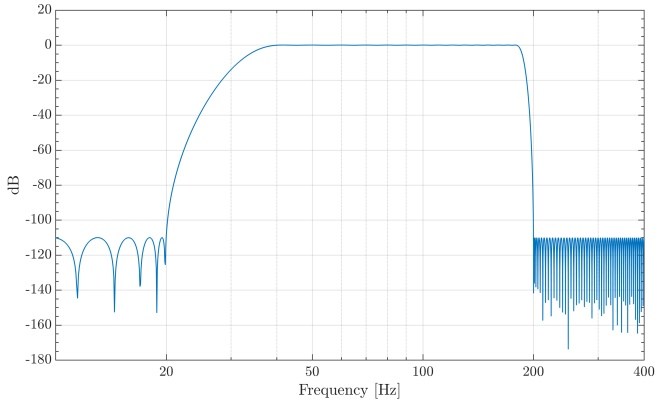


FIGURE 4.10: Bandpass filter for controlling target limits.

4.3.1.3 Regularisation

Regularisation was used in the calculation of an inverse filter to limit the amount of gain used to correct the transfer function when the determinant of the matrix to be inverted was getting close to zero. The inclusion of regularisation changed the Equation 4.10 from

$$\mathbf{h} = \left[\mathbf{C}^{\mathrm{H}}\mathbf{C}\right]^{-1}\mathbf{C}^{\mathrm{H}}\mathbf{a} \tag{4.37}$$

to

$$\mathbf{h} = \left[\mathbf{C}^{\mathrm{H}}\mathbf{C} + \beta\mathbf{I}\right]^{-1}\mathbf{C}^{\mathrm{H}}\mathbf{a}, \tag{4.38}$$

where βI is the Tikhonov regularisation parameter based on the work from Tokuno et al. (1997), Shin et al. (2014), Hansen (1998) and Hansen (2001). β was defined as

$$\beta(f) = ||\mathbf{C}(f)|| \times \eta, \tag{4.39}$$

where η is normally in the order of 10^{-2} to 10^{-4} (Hansen (1998)). For this work, an empirical method was used to set η, which is now considered to be not optimal to calculate the values of η. Although it did not degrade the result substantially, the L-curve method (Hansen (1998)) should have been used.

Figure 4.11 and 4.12 display the effect of regularisation on the designed filters $\eta = 10^{-2}$ and 10^{-3}, respectively. It was clear that the 10^{-2} had a larger effect on the filter design, so the decision was made to use the 10^{-3} value.

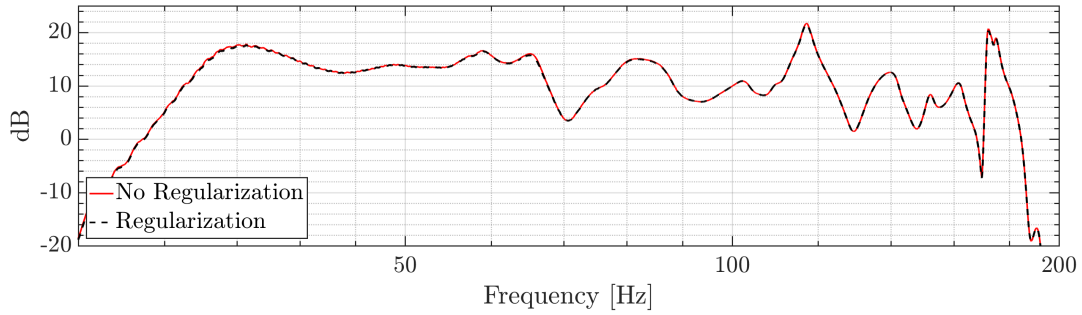


FIGURE 4.11: Filters designed with $\eta = 10^{-3}$.

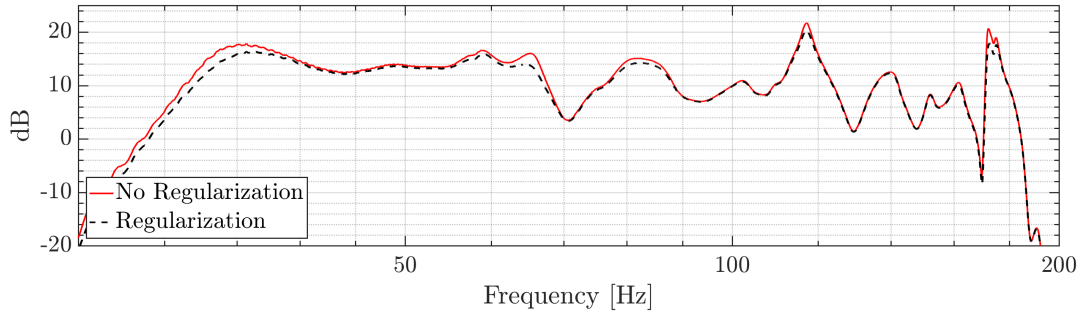
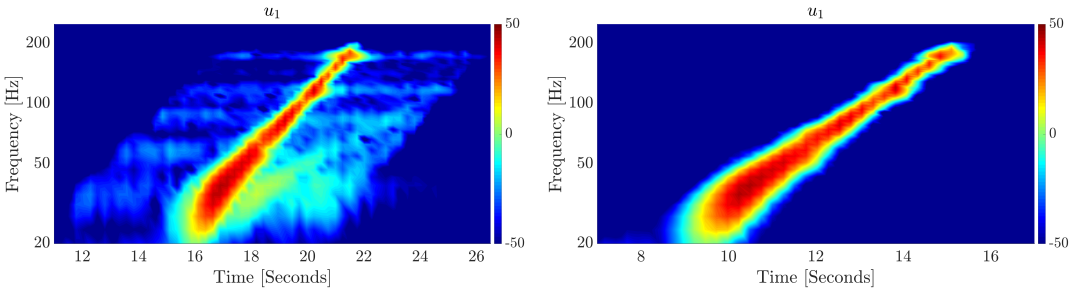


FIGURE 4.12: Filters designed with $\eta = 10^{-2}$.

4.3.1.4 Results from the improvements

The simplest way to get an idea of the changes that these improvements had made is to look at the filtered loudspeaker input signal (u) from **Step 5**. The spectrograms in Figure 4.13 plot the signal (u_1) without the improvements in Plot 4.13a and with

the improvements in Plot 4.13b. The change is clear to see, as the frequency smearing shown in the swept sine Plot 4.13a had been removed in the Plot 4.13b.



(a) Spectrogram of $(u)_1$ without improvements

(b) Spectrogram of $(u)_1$ with improvements

FIGURE 4.13: Spectrogram of the filtered swept sine for the plane wave target over four channels with and without the improvements.

4.3.2 Results of the real-world study

This section reports the real-world test data with all the outlined improvements. Two sets of data were compared. The first set of data replaced the Green's function transfer function in the analytical model with the real-world measurement to predict the subsequent real-world measurement. This was labelled as 4C $_{Model}$ and outlined in **Step 4**. The second applied the filters to the swept sine and took a second real-world measurement, and this was labelled as 4C $_{Real}$ and outlined in **Step 5**. Applying the low-frequency filter changes to the lower test data limits, the error was plotted from 20 to 100 Hz, but the frequency data were analysed from 30 to 100 Hz.

4.3.2.1 Plane wave target sound field

Figure 4.14 shows the error for all methods across the listening area for the frequency range of 20−100 Hz, and Table 4.5 shows the band-limited data for $\bar{\varepsilon}$, $|\text{RMS}_\Delta|$ from the target and the MSV.

Figure 4.14 shows the raw signal (—RS) at 0 dB and gives the reference point on which the other methods aim to improve. Moreover, it can be seen that both least-squares approaches (—4C $_{Model}$ & —4C $_{Real}$) were, at worst, 2 dB better than this line at 57.9 Hz. The real-world data tracked the model well above 45 Hz, with the largest delta of only 2 dB at approximately 65 Hz. Below 45 Hz, the real-world data did diverge from the model, but there was still approximately 15 dB of improvement over the raw signal. In Table 4.5, the $\bar{\varepsilon}$ data in this lower frequency band (30−40 Hz) of the real-world data shows a drop in effectiveness of 8.31 dB against the model.

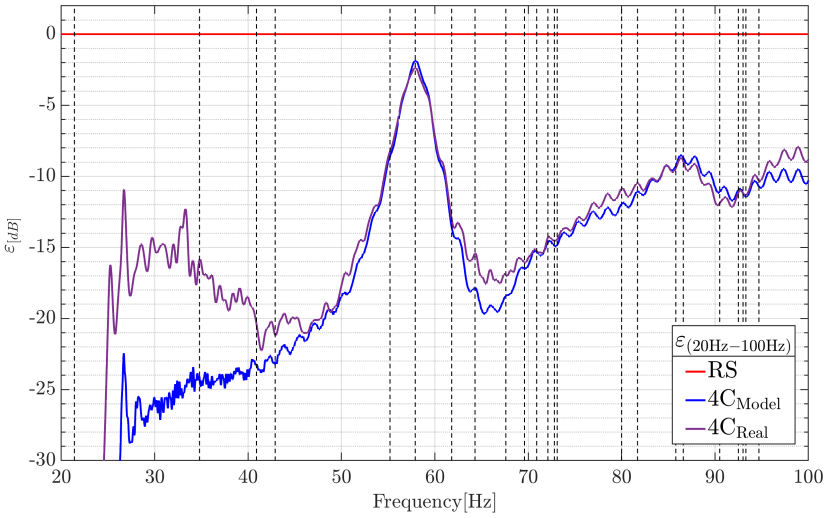


FIGURE 4.14: Error ratio ε against the frequency of the raw signal with no cancellation (—RS), of the four-channel approach model (—4C $_{Model}$) and four-channel real-world approach (—4C $_{real}$), over the fixed listening area with a plane wave target, with all the modal frequencies of the enclosure denoted as - -.

TABLE 4.5: Frequency banded mean error [dB], RMS [dB] and MSV [dB], over the fixed listening area with a plane wave target.

Method	30 Hz $> f \geq$ 100 Hz			30 Hz $> f \geq$ 40 Hz			40 Hz $> f \geq$ 70 Hz			70 Hz $> f \geq$ 100 Hz		
	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV
RS	0.00	1.93	35.49	0.00	2.44	41.96	0.00	2.03	28.56	0.00	3.04	40.13
$4C_{Model}$	-13.98	0.24	4.59	-25.40	0.18	0.09	-14.63	0.01	8.05	-12.34	0.59	2.65
$4C_{Real}$	-13.45	0.21	5.07	-17.09	0.14	0.56	-14.37	0.33	9.04	-12.04	0.10	2.64

In Table 4.5, the modelled data shows a good level of improvement from the raw data across the whole frequency band (30−100 Hz), with $\bar{\varepsilon}$ at -13.98 dB more effective than the raw signal and the real-world data within 0.53 dB to −13.45 dB. However, the data was less effective than the analytical model of the same room in the previous section. This was expected, as the model could not consider all environmental factors; nevertheless, this was still a good level of control. All the other metrics in Table 4.5 showed a good improvement on the raw signal, and the model data showed a good correlation with real-world data apart from the $|RMS_\Delta|$ in the upper-frequency range. Here, the model was outperformed by the real-world data and raw signal. The reason for this can be seen in Figure 4.15, which shows the SA across the listening area for the frequency range of 20−100 Hz with a plane wave target sound field. Above 70 Hz, both the model and the real-world data moved away from the target, but the real-world data stayed closer to the target and moved back to the target above 90 Hz.

In Figure 4.15, it is clear how well the model data could predict the real-world result. Apart from the large trough at 57.9 Hz, both sets of data were within 2 dB of the target

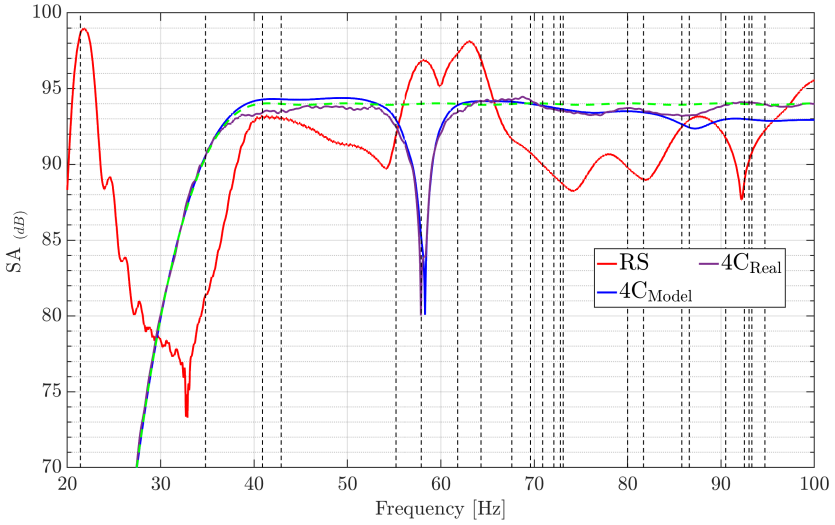


FIGURE 4.15: SA against the frequency of the raw signal with no cancellation (—RS), four-channel approach model (—4C $_{Model}$) and four channel real-world approach (—4C $_{real}$), over the fixed listening area with a plane wave target, with all the modal frequencies of the enclosure denoted as - -.

and showed a good level of control over the raw signal.

The first n_3 mode was 57.9 Hz, and it was also shown in Figure 4.14 as the least effective frequency of control. Figure 4.15 shows a relatively small peak in the raw signal SA at this frequency but a much larger trough in the SA for both the 4C $_{Model}$ and the 4C $_{Real}$, as if the filters built for the real-world measurements had overcompensated for the small peak. Through testing, it was inferred that this trough was due to a phase issue in the filters. This trough can no longer be seen if the filters are applied as zero-phase. However, changing to zero-phase filters creates a larger issue across the frequency domain. Therefore, it was accepted that the original filters should be used to get the best overall result.

In Table 4.5 the MSV data reinforced what was already asserted: both the model data and the real-world data showed a good level of control over the raw signal, and the model predicts the real world well. The MSV going from above the 35 dB unacceptable threshold for the raw signal, to below the 15 dB threshold for both the 4C $_{Model}$ and the 4C $_{Real}$, meaning that the four-channel least-squares approach channel demonstrates good real-world performance in improving the receiver-to-receiver variance.

4.3.2.2 Spherical wave target sound field

Moving from the plane wave target to the spherical wave target sound field, all the data in this section showed the same trends as in the previous sections, namely that both the

model data and the real-world data show a good level of control over the raw signal and that the model predicts the real world.

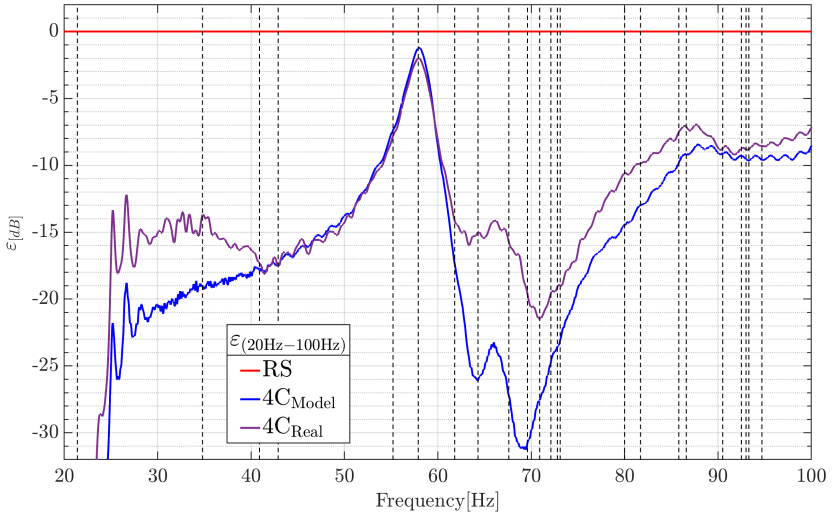


FIGURE 4.16: Error ratio ε against the frequency of the raw signal with no cancellation (—RS), the four-channel approach model (—4C $_{Model}$) and the four channel real-world approach (—4C $_{real}$), over the fixed listening area with a spherical wave target, with all the modal frequencies of the enclosure denoted as - -.

TABLE 4.6: Frequency banded mean error [dB], RMS [dB] and MSV over the fixed listening area with a spherical wave target.

Method	30 Hz $> f \geq$ 100 Hz			30 Hz $> f \geq$ 40 Hz			40 Hz $> f \geq$ 70 Hz			70 Hz $> f \geq$ 100 Hz		
	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV
RS	0.00	1.71	31.26	0.00	3.36	37.73	0.00	1.11	24.33	0.00	2.97	35.90
4C$_{Model}$	-13.74	0.23	2.06	-20.17	0.15	1.04	-13.41	0.04	5.22	-13.01	0.58	0.04
4C$_{Real}$	-12.62	0.15	2.33	-15.93	0.09	0.29	-13.42	0.32	5.18	-11.32	0.00	0.36

Figure 4.16 shows the raw signal (—RS) with both least-squares approaches (—4C $_{Model}$ & —4C $_{Real}$) and a sizable improvement of 12.62 dB for the real-world data and 13.74 dB for the model across the whole frequency band. Although the real-world data did not track the model as closely as in Figure 4.14, a delta of up to 10 dB could now be seen between the data sets in the range of 60−90 Hz. As discussed in the review of the plane wave result in the previous section, this is likely to be an issue with the phase in the filter. However, although the issue was clear in the range of 60−90 Hz in Figure 4.16, it was not reflected in the SA data shown in Figure 4.17, meaning the drop in effectiveness may not be audible to the end-user. In the band-limited data shown in Table 4.6, the modelled data was still a good predictor for all real-world metrics.

In Figure 4.17 (which shows the SA across the listening area for the frequency range of 20−100 Hz with a spherical wave target sound field), there is again a large trough at 57.9 Hz. However, both least-squares data sets again showed a good level of control

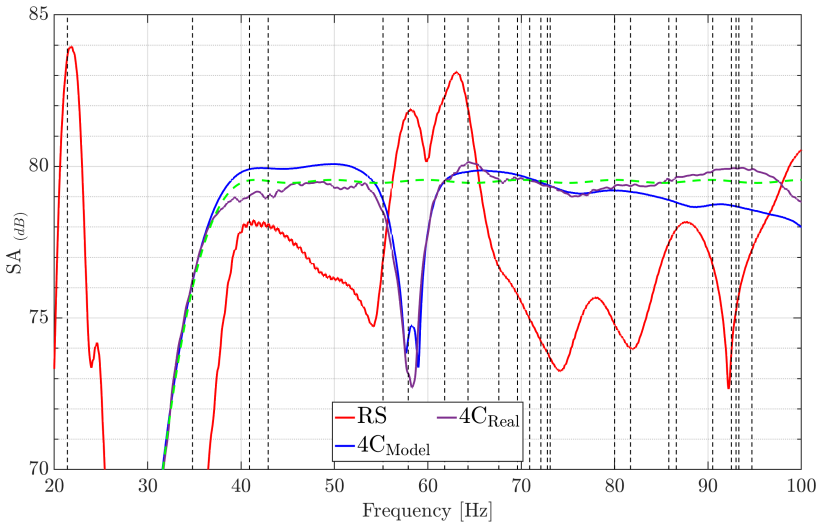


FIGURE 4.17: SA against the frequency of the raw signal with no cancellation (—RS), four-channel approach model (—4C $_{Model}$) and four channel real-world approach (—4C $_{real}$), over fixed listening area with a spherical wave target, with all the modal frequencies of the enclosure denoted as - -.

over the raw signal and were within 1 dB of the target up to 90 Hz. The model data showed a good level of prediction for the real-world data.

4.3.3 Study summary

This was a short real-world experimental study, and it was not long enough to draw major conclusions. However, there was enough data to make the following two statements:

1. Although less effective than the analytical model in Section 4.2, the level of control seen in the real-world data does show that the four-channel approach for acoustic control in cuboid rooms is effective when transferred to the real world.

2. When the transfer functions created using Green's functions are replaced with real-world data, the model offers a high level of accuracy in predicting real-world results.

4.4 Multiple room layouts

As real-world measurements can be time-consuming, the analytical model can be used to investigate multiple room layouts. This section provides the results from several

studies carried out to obtain a better understanding of the limitations and effectiveness of these four methods of control:

- two-channel (rear) least-squares approach
- three-channel least-squares approach
- four-channel least-squares approach
- plane wave cancellation method

Due to the amount of pre-existing work and the poor performance of the two-channel (front) control approach in Section 4.2, it was removed from the testing. However, the plane wave cancellation method was still included as a benchmark.

So far, all the studies using the analytical model except those in Section 3.3 used a layout that was optimal for the plane wave cancellation method. The aim of this section is to move away from the ideal layout in order to gain a better understanding of the performance of the least-squares approaches. The studies carried out for this section of the work are repetitive in nature, and the full details are in Appendix A. In this section, a brief overview of the setups and a summary of the findings are given.

4.4.1 Study setup

The dimensions of the rooms were again based on the Bowers & Wilkins listening room at their now closed SRE office, which has been used for all the models so far. In this section, the room was used in two orientations, as given in Table 4.7. These were the long thin room (LT) in the same orientation previously used and the short wide room (SW) (a rotation of 90 °).

For all studies, the analytical model was the same as in Section 4.2, which used Green's function to create transfer functions where $\zeta_{\mathbf{n}} = \frac{11.052}{\omega_n}$. The filters were created as described in Section 4.1, with both plane and spherical wave target sound fields used.

TABLE 4.7: Room orientation.

Name	l_1	l_2	l_3
Long Thin Room (LTR)	8 m	4.93 m	2.96 m
Short Wide Room (SWR)	4.93 m	8 m	2.96 m

Five layouts in total were studied in this work, namely

1. a layout ideal for the plane wave control method, denoted as layout C

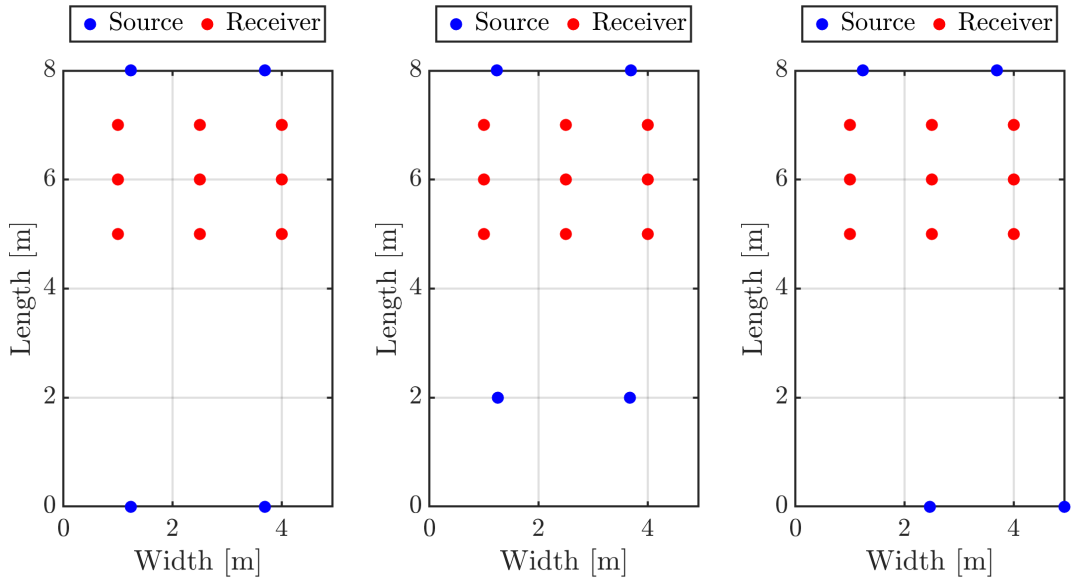
2. a layout where s_1 and s_2 were moved into a stereo placement, denoted as layout S
3. a layout where s_1 and s_2 were offset, denoted as layout O
4. a layout with dual listening areas, denoted as layout D
5. a layout with triple listening areas, denoted as layout T

For all studies in the long thin room, the receiver locations are given in Table 3.6. Three source configurations were studied, the locations of which are listed in Table A.2. As shown in the graphical representation in Figure 4.18, they all focused on moving the front two sources s_1 and s_2. These layouts were

Layout 1: **LTC** (Plot 4.18a) - ideal for the plane wave control method up to 69.6 Hz, as calculated using (3.7).

Layout 2: **LTS** (Plot 4.18b) - s_1 and s_2 were moved into a stereo placement.

Layout 3: **LTO** (Plot 4.18c) - s_1 and s_2 were offset.



(a) LTC layout. (b) LTS layout. (c) LTO layout..

FIGURE 4.18: Graphical representations of the LT room layouts in the test.

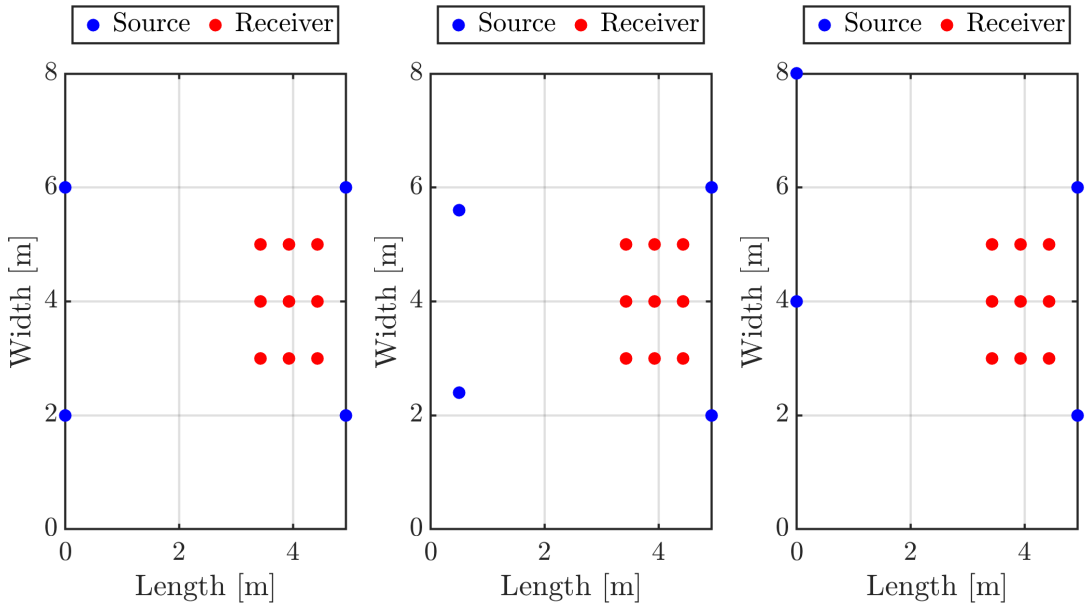
There were two reasons for the short wide room studies. This first reason was to understand how effective these systems can be when l_1 is not the largest dimension in the environment. The second reason was that the dimensions of the short wide room allowed more listening areas. Each study had three sets of results, and for all studies, the receiver locations when one listening area was in use are given in Table A.4. In the

first short wide room study, three source configurations were used, and the locations are given in Table A.5. These are shown in the graphical representation in Figure 4.19. They all focused on moving the two front sources, s_1 and s_2. The layouts were

Layout 1: **SWC** (Plot 4.19a) - an ideal layout for plane wave control. However, the increased distance between the sources should lower the maximum frequency of a plane wave that can be created (as outlined in Section 3.1) to 42.9 Hz, as calculated using (3.7).

Layout 2: **SWS** (Plot 4.19b) - s_1 and s_2 were moved into a stereo placement.

Layout 3: **SWO** (Plot 4.19c) - s_1 and s_2 were offset.



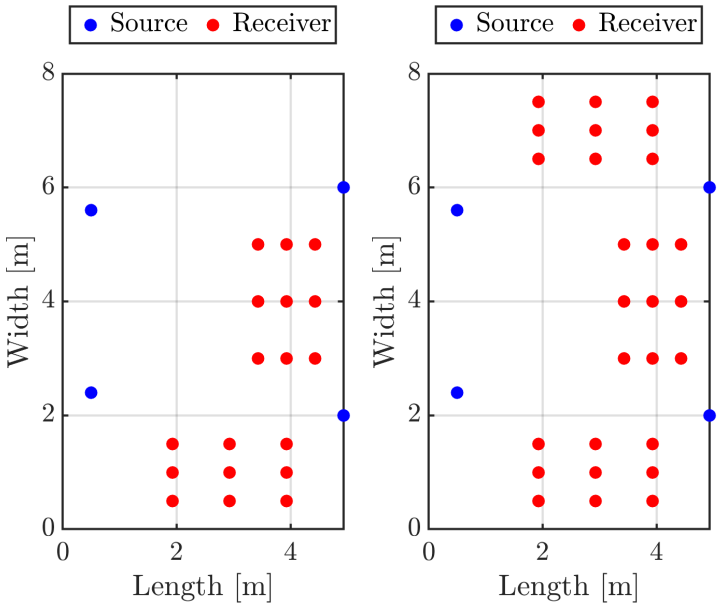
(a) SWC. (b) SWS. (c) SWO.

FIGURE 4.19: SW room layouts.

Previous studies looked at the effectiveness of the control over a single listening area while changing the front source locations. Here, the stereo layout was used, but the effect of the added second and third listening areas was investigated. The locations of the sources are given in Table A.5 as SWS, and the first listening zone receiver locations are given in Table A.4, the second in Table A.7 and third in Table A.8. Figure 4.20 shows a graphical representation of the two layouts in the test. These were compared with the SWS data from the previous section:

Layout 1: **SWD** (Plot 4.20a) - dual listening areas. These also showed the effectiveness of an asymmetrical layout.

Layout 2: **SWT** (Plot 4.20b) - triple listening areas.



(a) SWD (b) SWT.

FIGURE 4.20: SW room listening zone layouts.

4.4.2 Study summary

These studies continued to expand the information on the effectiveness and possible limitations of the least-squares approaches.

- The four-channel, three-channel and two-channel (rear) approaches were always, on average, more effective than the plane wave cancellation method. Although a change in layout changed the effectiveness of these methods, the decline in effectiveness was deemed small enough not to be a limitation of the least-squares approaches.
- Only the two-channel (rear) approach showed the limitation of being worse than the raw signal when using a long thin room in a stereo layout with a plane wave target.
- When l_1 was not the longest dimension in the environment, all the methods were less effective, but the drop in the effectiveness of the four-channel approach was less than the drop in the effectiveness of the plane wave cancellation method.
- The effectiveness of the least-squares approaches decreased as the number of areas of control increased, whereas on the whole, the plane wave cancellation method was constant, even though the least-squares approaches still outperformed the plane wave cancellation method. This was to be expected. As discussed in Chapter 3, the plane wave cancellation method was developed to control the sound field across a whole room. Therefore, its performance should be constant

no matter how many listening zones are used. However, the effectiveness of the least-squares approaches is inversely proportional to the number of listening zones used. Therefore, the more zones that are in use, the less targeted the least-squares approaches control can be for any zone-specific issues. Nevertheless, it is theorised that when trying to control the sound field across a whole room, the four-channel approach should still be no worse than the plane wave cancellation method, which is investigated later in this chapter.

The next discussion focuses only on the findings from the plane wave cancellation method and four-channel approach. Figure 4.21 and Figure 4.22 cross-compares the band-limited $\bar{\varepsilon}$ values for both room orientations with the plane and spherical wave target sound fields.

In both Figure 4.21 and Figure 4.22 it is shown that layout C was the most effective in all data sets. The reason for this (covered in Section 3.3.2) is that when a source is on the boundary, the signal from that source propagates as a single wavefront. However, when a source is just off the boundary, the signal from that source propagates as a wavefront, closely followed by the delayed reflection from the boundary. It is always going to be easier to control a single wavefront. The drop in the effectiveness of layouts C and O between the two room orientations is interesting. The change in the distance between sources s_1 and s_2 from the LT room to the SW room decreased the maximum frequency at which the image method (introduced in Section 3) could create a plane wave. The effect of this was clear in the data for the plane wave cancellation method in Figure 4.21 in the $70-100$ Hz band range for layout C. However, this did not explain the delta in the plane wave cancellation method shown in the same figure in the range of 10–40 Hz. This range was below the maximum frequency at which a plane wave could be created for both orientations. It is thought that this is due to the distance between the sources and the location of the listening zone. In the LT room, the distance between sources (or imaged sources) was 2.47 m with the listening zone starting 5 m away. In the SW Room, the distance was 4 m with the listening zone starting 3.43 m away. This meant that in the SWR, the listening zone started in the near field to the sources before the image method formed the plane wave.

In Section 3.3.2, it was shown that the largest drop in the effectiveness of the plane wave cancellation method was as the front sources were moved away from the front wall. The data here indicates that the same may be true for the four-channel approach, which is investigated later in this chapter. The data also shows that the drop in effectiveness when moving the front sources away from the front wall may be inversely proportional to the room's width and the distance between the sources.

All gathered data support the use of the four-channel least-squares approach. For the rest of this thesis, only the four-channel least-squares approach is used, as it is clearly the most effective in all given situations.

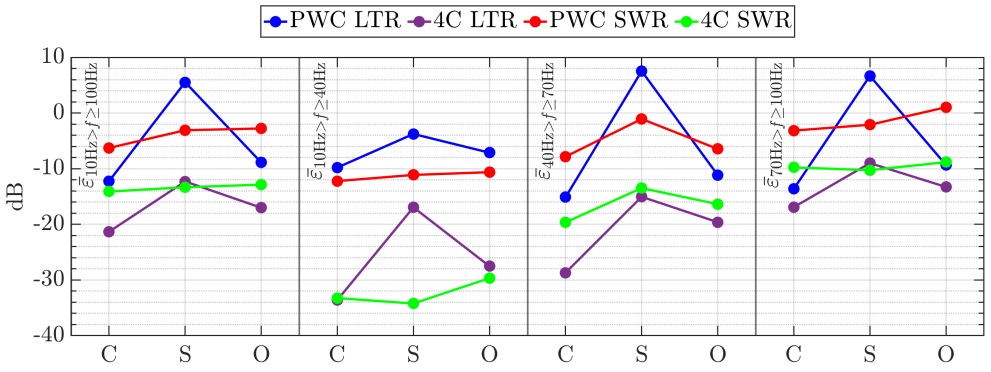


FIGURE 4.21: Compares the band-limited $\bar{\varepsilon}$ [dB] over the fixed listening area between the LT and SW Rooms, with the C, S, and O layouts with plane wave target sound fields, for the plane wave cancellation method (—●—PWC LT room), the four-channel approach (—●—4C LT room), the plane wave cancellation method (—●—PWC SW room), and the four-channel approach (—●—4C SW room).

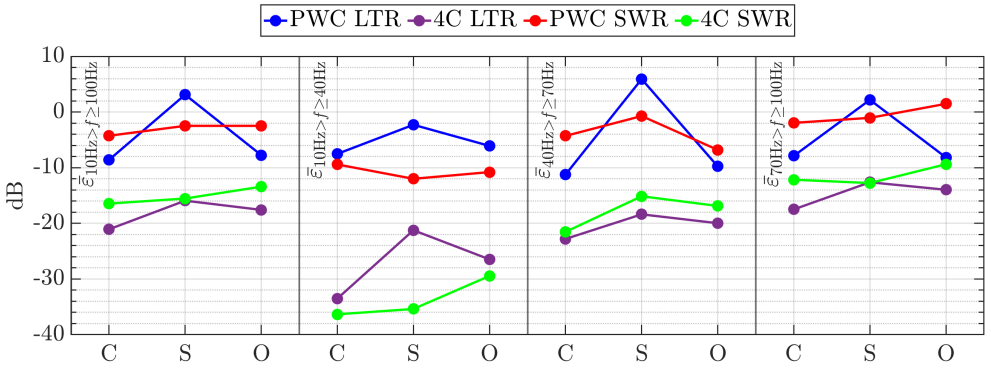


FIGURE 4.22: Compares the band-limited $\bar{\varepsilon}$ [dB] over the fixed listening area between the LT and SW rooms, for the C, S, and O layouts with spherical wave target sound fields, of the plane wave cancellation method (—●—PWC LT room), the four-channel approach (—●—4C LT room), the plane wave cancellation method (—●—PWC SW room), and the four-channel approach (—●—4C SW room).

4.5 Limitations of the four-channel approach

The four-channel least-squares approach was shown in the previous sections to be the most effective of all the methods studied here in controlling the sound field of the listening area(s). The aim of this section is to further compare the four-channel least-squares approach with the plane wave cancellation method to gain more information about the system's limitations. In order to do this, a number of studies were carried out.

- The three studies from Section 3.3 were repeated with the addition of the four-channel least-squares approach.

- A comparison between the plane wave cancellation method and the four-channel least-squares approach over a whole enclosure rather than a fixed listening area was made.
- A comparison of 3,000 source location combinations over a fixed listening area was made. This was not about finding the optimal source locations but investigating how the source locations alter the effectiveness of the control method.

Unless otherwise stated in the following studies, a long thin room as defined in Table 4.7 was used with a stereo layout, the source and receiver locations as given in Table A.2 and Table 3.6, respectively. The analytical model is defined in Section 4.2, and the damping of $\zeta_{\mathbf{n}} = \frac{11.052}{\omega_n}$.

4.5.1 Rerunning the limitations of the plane wave cancellation method test from chapter 3

In this section, the studies from Section 3.3 were repeated with the plane wave cancellation method now compared with the four-channel least-squares approach.

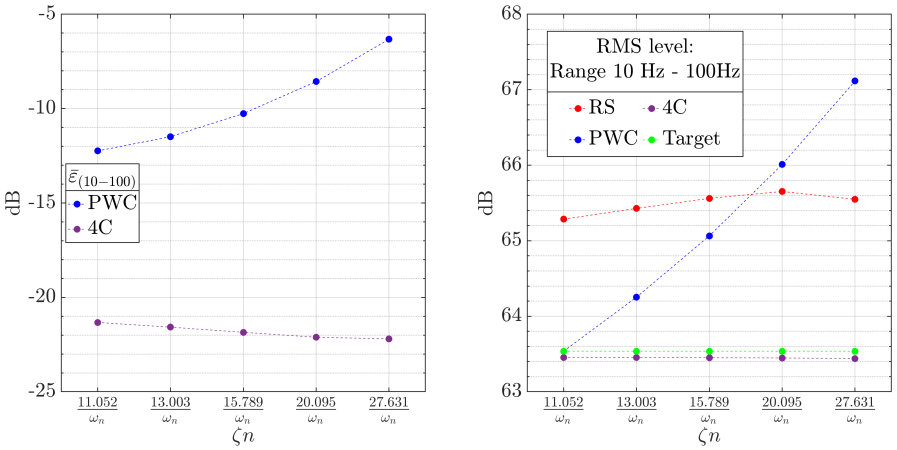
4.5.1.1 The effect of damping

As in Section 3.3.1, this study compared the results from five different damping coefficients namely when $\zeta_{\mathbf{n}} = \frac{11.052}{\omega_n}$, $\zeta_{\mathbf{n}} = \frac{13.003}{\omega_n}$, $\zeta_{\mathbf{n}} = \frac{15.789}{\omega_n}$, $\zeta_{\mathbf{n}} = \frac{20.095}{\omega_n}$ and $\zeta_{\mathbf{n}} = \frac{27.631}{\omega_n}$.

Figure 4.23 shows the changes in $\bar{\varepsilon}$ (Plot 4.23a), the RMS level (Plot 4.23b) and the MSV (Plot 4.23c) across the frequency range as the damping of the enclosure was increased. In Plot 4.23a, it is clear that the four-channel least-squares approach was not just more effective than the plane wave cancellation method for all damping coefficients tested, but unlike the plane wave cancellation, the four-channel approach became slightly (approximately 0.75 dB) more effective as the damping increased. This is due to the fact that as damping increases, the environment is less reverberant and therefore easier for the four-channel approach to control.

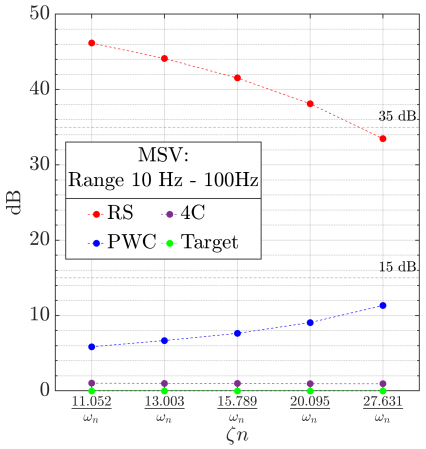
When $\zeta_{\mathbf{n}} = \frac{11.052}{\omega_n}$, it was expected that the plane wave cancellation method RMS level in Plot 4.23b would match the target (as explained in Chapter 3). The four-channel approach was never the same as the target but tracked the target at a –0.17 dB difference at each point, and it was more effective than the plane wave cancellation method once the damping increased.

In Section 3.3.1, it was discussed that if the error in the plane wave cancellation method, as the damping of the enclosure increases, is only due to the gain change, it can be expected that the plane wave cancellation method MSV would show a flatter line. Even if



(a) Change in $\bar{\varepsilon}$.

(b) Change in RMS.



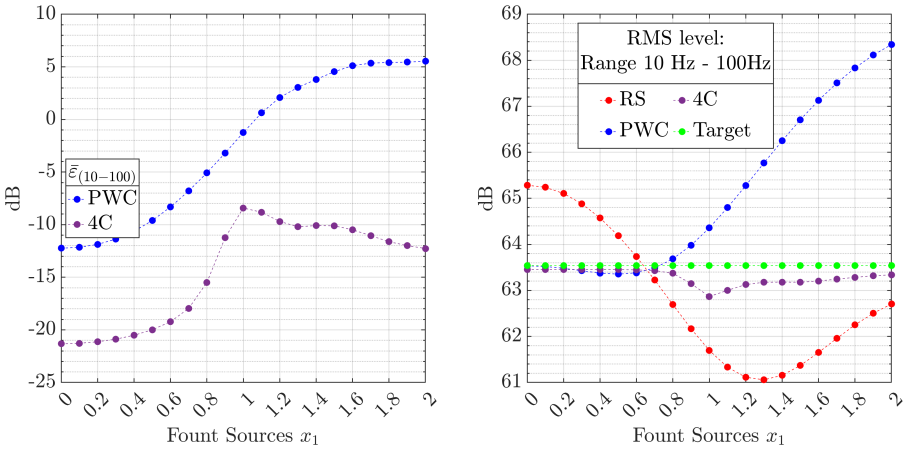
(c) Change in MSV.

FIGURE 4.23: $\bar{\varepsilon}$, RMS and MSV for different damping coefficients. The raw signal (-●-RS), the plane wave cancellation method (-●-PWC), the four-channel approach (-●-4C) and the target (-●-Target) are shown.

the level changes at each receiver point, the variance should stay the same. This can be seen in the four-channel approach in Plot 4.23c, meaning not only did the four-channel approach outperform the plane wave cancellation method, but the performance was more consistent. Although both the plane wave cancellation method and the four-channel approach are under the 15 dB threshold that denotes a good system, the four-channel approach shows a variance of approximately 1 dB for all damping values.

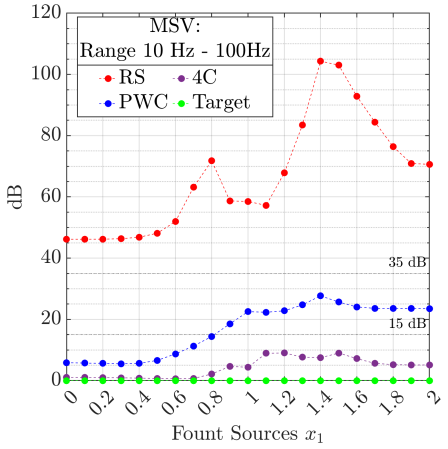
4.5.1.2 The effect of moving s_1 and s_2 perpendicular to the front wall

The study in Section 3.3.2 was repeated, wherein the sources s_1 and s_2 were moved as a pair from the ideal position on the front wall into the room in 0.1 m increments along the x-axis from 0 to 2 m to provide 21 sets of data. When using the plane wave cancellation method, as the front sources were moved into the room, the time delay of the rear source signals was lowered to compensate for the new locations.



(a) Change in $\bar{\varepsilon}$

(b) Change in RMS.



(c) Change in MSV.

FIGURE 4.24: Change in $\bar{\varepsilon}$, RMS and MSV when moving the front sources perpendicular to the front wall. The raw signal (-●-RS), the plane wave cancellation method (-●-PWC), the four-channel approach (-●-4C) and target (-●-Target) are shown.

Plot 4.24a shows the $\bar{\varepsilon}$ over the frequency range. At all points, the four-channel approach outperformed the plane wave cancellation method. However, the effectiveness of the four-channel approach decreased as the sources were moved away from the front wall. At 1 m, the four-channel approach had a minimum effectiveness of approximately –8.25 dB, which did show a limitation of the approach. However, at the same location, the plane wave cancellation method was approximately -1 dB. This minimum effectiveness of the four-channel approach still improved the sound field.

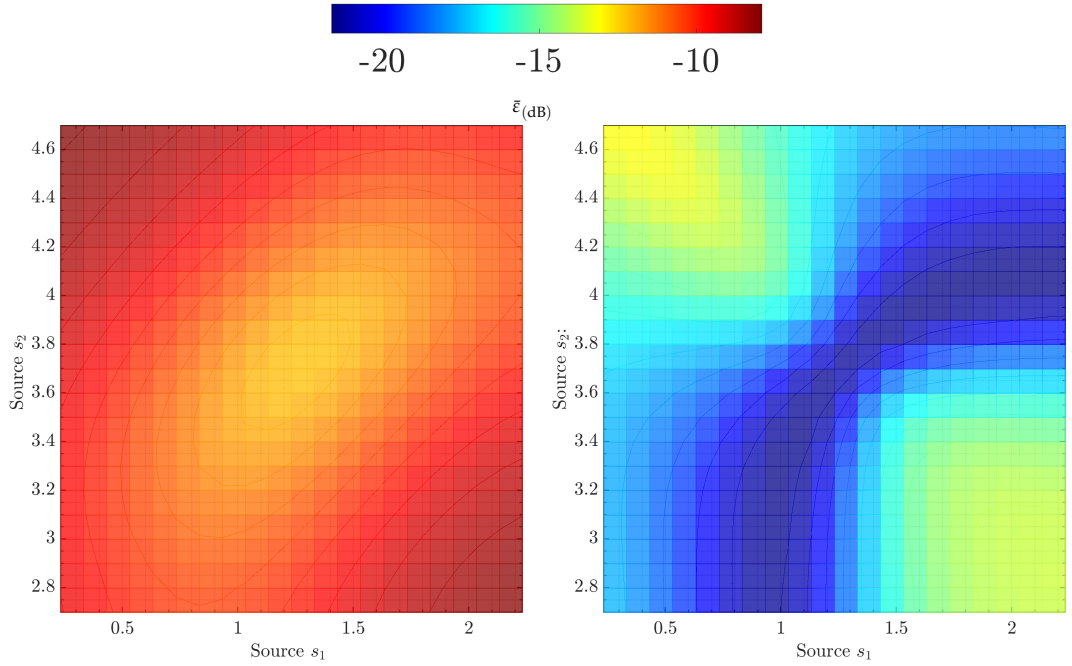
Plot 4.24b shows the change in the RMS level over the frequency range as the source location changed. As expected, the plane wave cancellation method RMS level matched the target when the sources were in the ideal locations. Up to 0.2 m and between 0.7 m and 0.9 m, the plane wave cancellation method outperformed the four-channel approach. The four-channel approach was never closer than –0.17 dB from the target. However, once the plane wave cancellation method started to diverge from the target,

the four-channel approach was clearly better because, at worst, it was only approximately 0.6 dB from the target.

In Figure 4.24c, the four-channel approach was lower in value than the plane wave cancellation method. Although the MSV does rise as the sources move away from the front wall, the four-channel approach is always below the 15dB threshold. However, the plane wave cancellation method goes over the 15 dB threshold at 0.9 m and is only in the acceptable range for subsequent data sets.

4.5.1.3 The effect of moving s_1 and s_2 parallel to the front wall

The study in Section 3.3.3 was repeated, wherein the sources s_1 and s_2 were fixed on the front wall but moved along the y-axis. Each source was moved independently up to $\pm$ 1.0 m (in 0.1 m increments) from their ideal positions along the same horizontal plane, resulting in 421 permutations for the pair of front sources.



(a) The plane wave cancellation method.

(b) Four-channel least-squares approach.

FIGURE 4.25: Surface plot of $\bar{\varepsilon}$ for 421 pairs of the sources s_1 and s_2. The locations of the front sources are along the front wall.

Figure 4.25 shows the $\bar{\varepsilon}$ values for 421 pairs of sources s_1 and s_2. Their locations using the plane wave cancellation method are shown in Plot 4.25a and the four-channel least-squares approach in Plot 4.25b. From these two plots, it is clear that the four-channel approach was more effective for all source locations. These data sets are clearer in the histogram shown in Figure 4.26, which shows that the difference between the least

effective values of the four-channel approach and the most effective values of the plane wave cancellation method was over 0.5 dB. The range of the four-channel approach was wider, which meant that the effectiveness was more susceptible to the locations of the sources, s_1 and s_2. Figure 4.26 introduces a new type of plot to the thesis, a histogram, to allow easier comparison between the two sets of data. Each set of data had the same bin size and was plotted as a normalised probability density function (pdf).

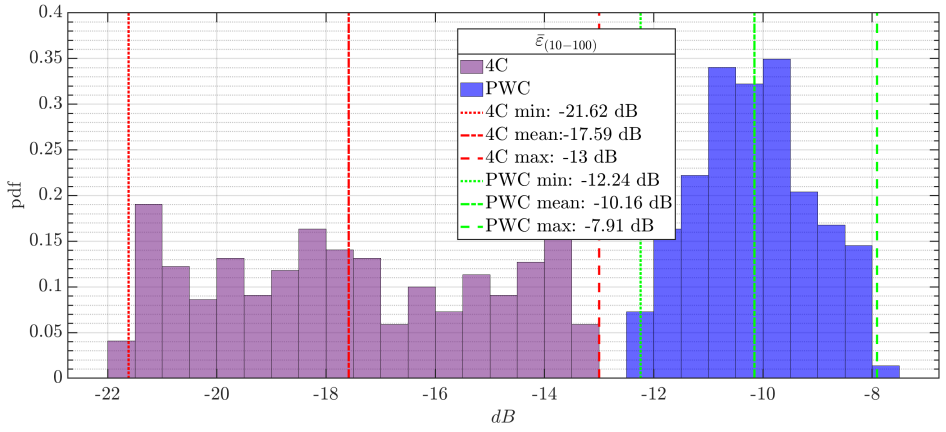
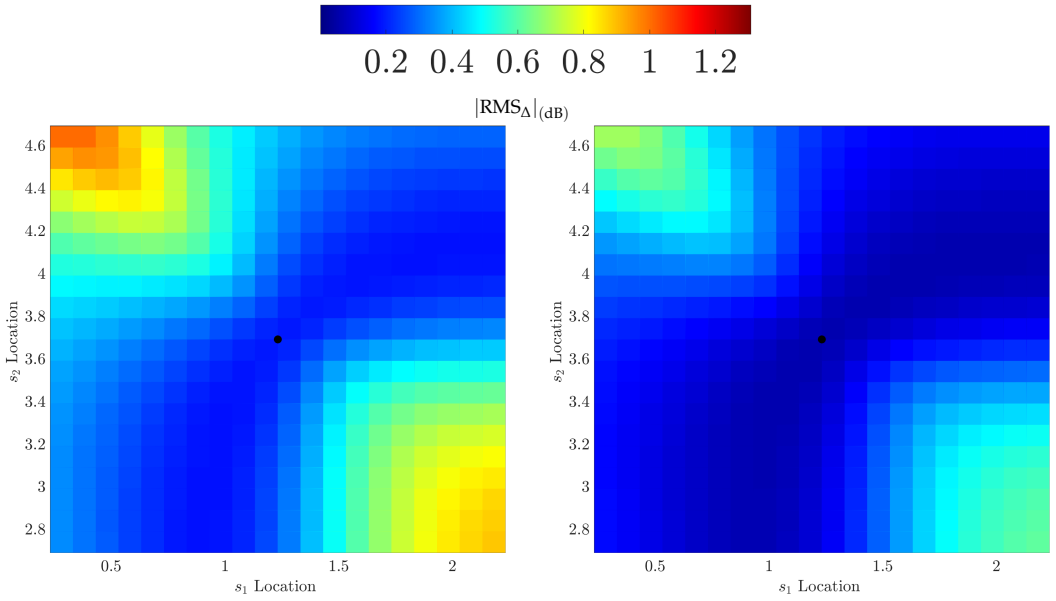


FIGURE 4.26: Histograms of $\bar{\varepsilon}$ for 421 pairs of the sources s_1 and s_2. The locations of the front sources are along the front wall for the plane wave cancellation method (PWC) and four-channel approach (4C).



(a) The plane wave cancellation method.

(b) Four-channel least-squares approach.

FIGURE 4.27: Surface plot of $|\text{RMS}_\Delta|$ for 421 pairs of the sources s_1 and s_2. The locations of the front sources are along the front wall.

Figure 4.27 shows the $|\text{RMS}_\Delta|$ from the target for 421 pairs of the sources s_1 and s_2 using the plane wave cancellation method in Plot 4.27a and the four-channel least-squares

approach in Plot 4.27b. In both plots, the lowest values of $|\mathrm{RMS}_\Delta|$ from the target indicated by the black dots (the ideal positions of s_1 and s_2) were 0 dB for the plane wave cancellation method and 0.08 dB for the four-channel approach. In Figure 4.27, the plane wave cancellation method appears more effective at achieving the RMS target for a wider range of source locations (also shown in Figure 4.28). However, Figure 4.28 also shows that the four-channel approach had a smaller range of 0.86 dB versus 1.75 dB and that both methods had a similar mean. Although the four-channel approach did not obtain the target RMS over the listening area, its achievable RMS level was less susceptible to the locations of the sources s_1 and s_2.

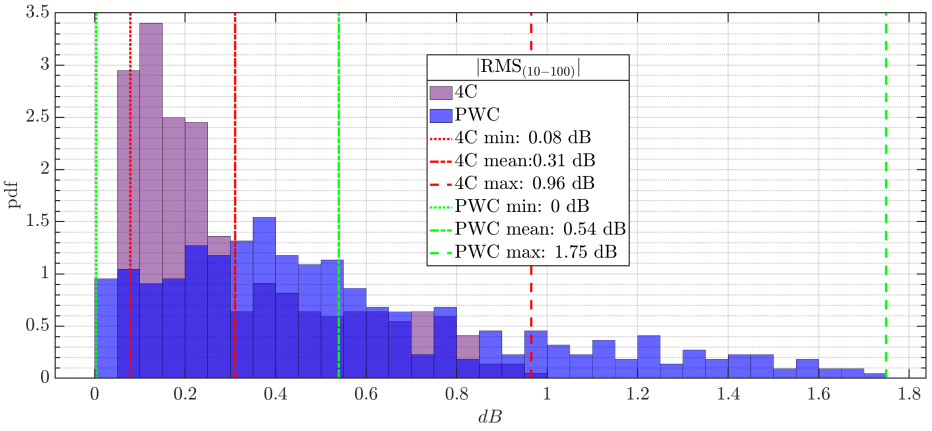
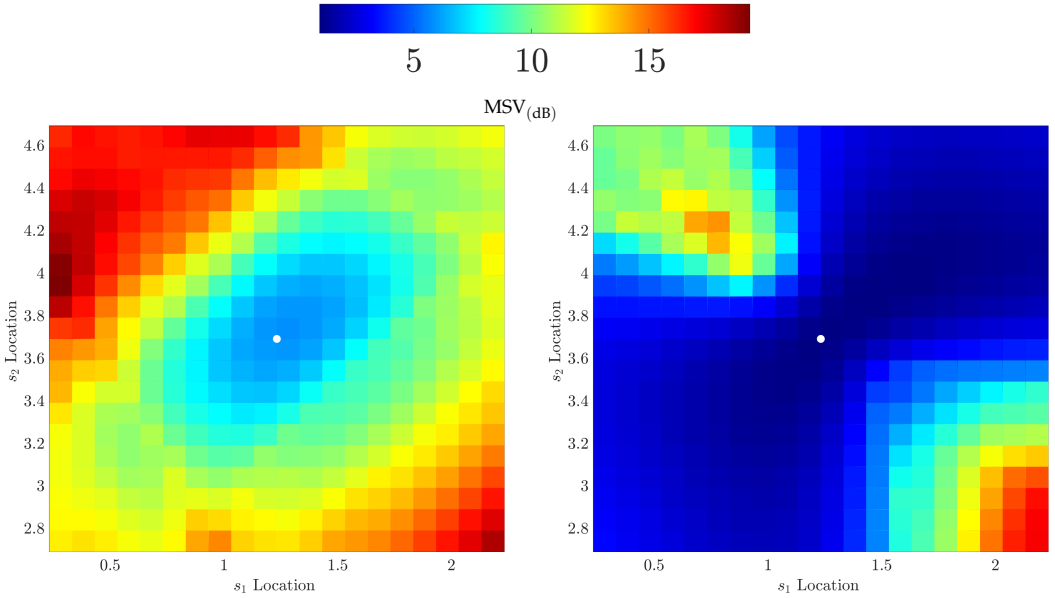


FIGURE 4.28: Histograms of $|\mathrm{RMS}_\Delta|$ for 421 pairs of the sources s_1 and s_2 with the locations of the front sources along the front wall, for the plane wave cancellation method (PWC) and the four-channel approach (4C).

Figure 4.29 shows the MSV from the target with 421 permutations for the pair of front sources. The locations using the plane wave cancellation method are shown in Plot 4.29a and the four-channel approach in Plot 4.29b. In both plots, the lowest values of MSV from the target indicated by the white dots (at the ideal positions of s_1 and s_2) were 5.85 dB for the plane wave cancellation method and 0.99 dB for the four-channel approach. In Figure 4.29b, the only drop from a good level of variation to an acceptable level of variation was when the sources were at their closest. This may be a limitation in that there must be a minimum distance between the sources for the four-channel approach to be effective. However, as shown in Figure 4.30, the mean of the four-channel approach outperformed the plane wave cancellation method by 7.23 dB.

4.5.1.4 Study summary

Repeating the studies from Section 3.3 comparing the plane wave cancellation method with the four-channel approach and for the cases tested here, the following results were obtained:



(a) The plane wave cancellation method.

(b) Four-channel least-squares approach.

FIGURE 4.29: Surface plot of MSV for 421 pairs of the sources s_1 and s_2 with the locations of the front sources along the front wall.

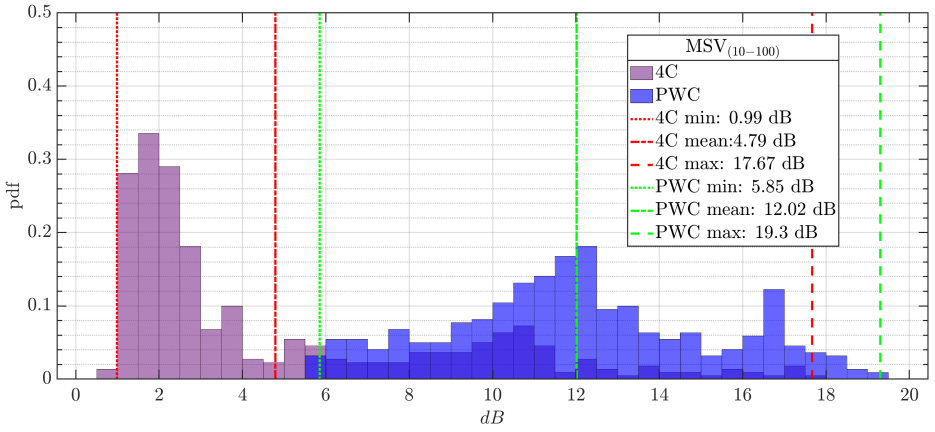


FIGURE 4.30: Histograms of MSV for 421 pairs of the sources s_1 and s_2, with the locations of the front sources along the front wall, for the plane wave cancellation method (PWC) and the four-channel approach (4C).

- In the vast majority of the cases, the four-channel approach outperformed the plane wave cancellation method. The only metrics that show this is not the case is where the RMS values of the plane wave cancellation method achieved the target when used in an ideal setup and $\zeta_{\mathbf{n}} = \frac{11.052}{\omega_n}$. However, as discussed in Chapter 3, the RMS level of the plane wave target was chosen specifically as a value that can be achieved by the plane wave cancellation method in this setup and when $\zeta_{\mathbf{n}} = \frac{11.052}{\omega_n}$.

- Changing the damping coefficients resulted in little change in the effectiveness of the four-channel approach.
- There was a degradation in the effectiveness of the four-channel approach when the front sources were moved away from the wall. However, the four-channel approach always showed an improvement over the raw signal.
- The four-channel approach $\bar{\varepsilon}$ was more susceptible than the plane wave cancellation method to the location of the source on the wall but always outperformed it.

4.5.2 Whole room

As discussed in Chapter 3, the plane wave cancellation method controls the whole room, and in Section 4.4.2 it was shown that as the number of listening areas increases, the effectiveness of the four-channel approach decreases. This study investigates whether the four-channel approach is limited to the listening areas or can also control the whole room. The modelled room was split into a grid with a receiver point at every 0.1 m in the x direction and 0.0986 m in the y. This grid was a flat surface with a height of 1.48 m. In this study, the target sound field across the whole room was a plane wave, as defined in Equation 4.34.

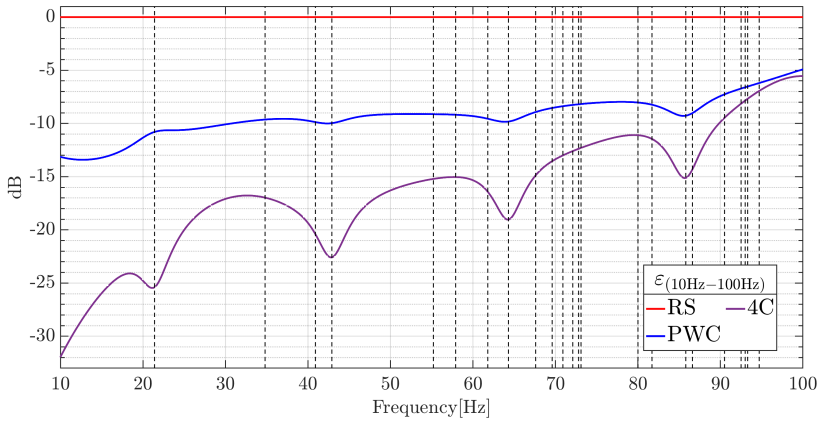


FIGURE 4.31: Error ratio ε against the frequency of the raw signal with no cancellation (—RS), the plane wave cancellation method (—PWC) and four-channel approach (—4C), over the whole room with a plane wave target, with all the modal frequencies of the enclosure denoted as - -.

Figure 4.31 shows the results of the error ratio for the raw signal (—RS), plane wave cancellation method (—PWC) and four-channel approach (—4C) across the defined surface in the room. The figure shows that the four-channel approach outperformed the plane wave cancellation method in achieving a plane wave over the room. As theorised

in Section 4.1.2.3, the four-channel approach can never be worse than the plane wave cancellation method due to the degree of freedom in the calculation of the filters.

TABLE 4.8: Frequency banded mean error [dB], RMS [dB] and MSV over the whole room with a plane wave target.

Method	30 Hz $> f \geq$ 100 Hz			30 Hz $> f \geq$ 40 Hz			40 Hz $> f \geq$ 70 Hz			70 Hz $> f \geq$ 100 Hz		
	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV
RS	0	1.81	29.57	0	1.99	25.37	0	0.85	30.59	0	2.82	32.82
PWC	-8.98	4.02	8.01	-10.86	2.3	5.17	-9.33	5.45	8.97	-7.43	5.04	9.91
4C	-13.26	0.46	2.46	-19.84	0.1	0.84	-16.52	0.25	1.06	-9.62	1.09	5.48

Table 4.8 compares the data metrics for the plane wave target sound field across the whole room. It shows that the four-channel approach was less effective when targeting the whole room (compared to a single listening area in Section 4.2.1.1), but still offered a good level of improvement over the raw signal and the plane wave cancellation method. Therefore, the four-channel approach does not have to be limited to a fixed listening area.

4.5.3 Testing the source location

In this study, there were 25 possible front source locations, shown in Table 4.9, and five possible rear source locations, shown in Table 4.10, with the receiver locations given in Table 3.6. This layout is represented in Figure 4.32. All possible combinations of the source locations were tested following the rules that only s_1 & s_2 can be the front source locations and only s_3 & s_4 can be the rear source locations.

The motivation behind this study was to understand more about the possible effect of an end-user placing the source at any location in the listening environment. It has been discussed how the plane wave cancellation method is constrained by source location. Therefore, when using the plane wave cancellation method, as the front sources were moved into the room, the time delay of the rear source signals was lowered to compensate for the new locations. It was impractical to test all possible source locations, but this study aimed to highlight any constraint to the four-channel approach due to the source locations.

TABLE 4.9: Possible front source locations.

Source No.	Source Location ($\bar{x}$ m, $\bar{y}$ m, $\bar{z}$ m)	Source No.	Source Location ($\bar{x}$ m, $\bar{y}$ m, $\bar{z}$ m)	Source No.	Source Location ($\bar{x}$ m, $\bar{y}$ m, $\bar{z}$ m)	Source No.	Source Location ($\bar{x}$ m, $\bar{y}$ m, $\bar{z}$ m)	Source No.	Source Location ($\bar{x}$ m, $\bar{y}$ m, $\bar{z}$ m)
1	(0.50, 0.50, 1.48)	6	(1.50, 0.50, 1.48)	11	(2.50, 0.50, 1.48)	16	(3.50, 0.50, 1.48)	21	(4.50, 0.50, 1.48)
2	(0.50, 1.4825, 1.48)	7	(1.50, 1.4825, 1.48)	12	(2.50, 1.4825, 1.48)	17	(3.50, 1.4825, 1.48)	22	(4.50, 1.4825, 1.48)
3	(0.50, 2.465, 1.48)	8	(1.50, 2.465, 1.48)	13	(2.50, 2.465, 1.48)	18	(3.50, 2.465, 1.48)	23	(4.50, 2.465, 1.48)
4	(0.50, 3.4475, 1.48)	9	(1.50, 3.4475, 1.48)	14	(2.50, 3.4475, 1.48)	19	(3.50, 3.4475, 1.48)	24	(4.50, 3.4475, 1.48)
5	(0.50, 4.4300, 1.48)	10	(1.50, 4.4300, 1.48)	15	(2.50, 4.4300, 1.48)	20	(3.50, 4.4300, 1.48)	25	(4.50, 4.4300, 1.48)

TABLE 4.10: Possible rear source locations.

Source No.	Source Location ($\bar{x}$ m, $\bar{y}$ m, $\bar{z}$ m)
26	(7.50, 0.50, 1.48)
27	(7.50, 1.4825, 1.48)
28	(7.50, 2.465, 1.48)
29	(7.50, 3.4475, 1.48)
30	(7.50, 4.43, 1.48)

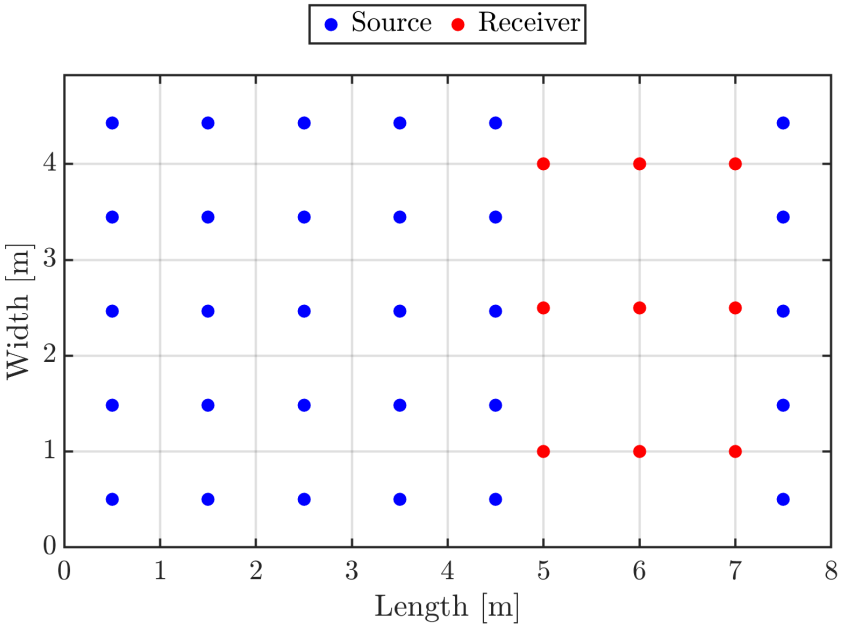


FIGURE 4.32: All possible source locations.

The target sound field over the listening area in this study was a plane wave, as defined in Equation 4.34. The data sets are listed in Figures 4.33, 4.34 and 4.35. Each figure shows two histograms for the plane wave target when using all defined source locations for both the four-channel approach and the plane wave cancellation method, highlighting the minimum, maximum and average values for each set of data.

In Figure 4.33, the two histograms show the $\bar{\varepsilon}$ over the frequency band. The effectiveness of the four-channel approach clearly relied on the source locations, as the results show a range of effectiveness varying by 10.27 dB from –16.37 dB to –6.1 dB. However, $\bar{\varepsilon}$ data show that, across the full frequency band, the four-channel approach still offered an improvement over the raw signal at the 0 dB reference at all the locations and a mean effectiveness of –9.7 dB. Comparing the four-channel approach with the plane wave cancellation method, the plane wave had a range of 18.88 dB, and the vast majority of the source locations provided a $\bar{\varepsilon}$ that offered a worse outcome than the raw signal at the 0 dB reference. It was outperformed by the four-channel approach.

In Figure 4.34, the two histograms show the $|\mathrm{RMS}_\Delta|$ from the target in the frequency

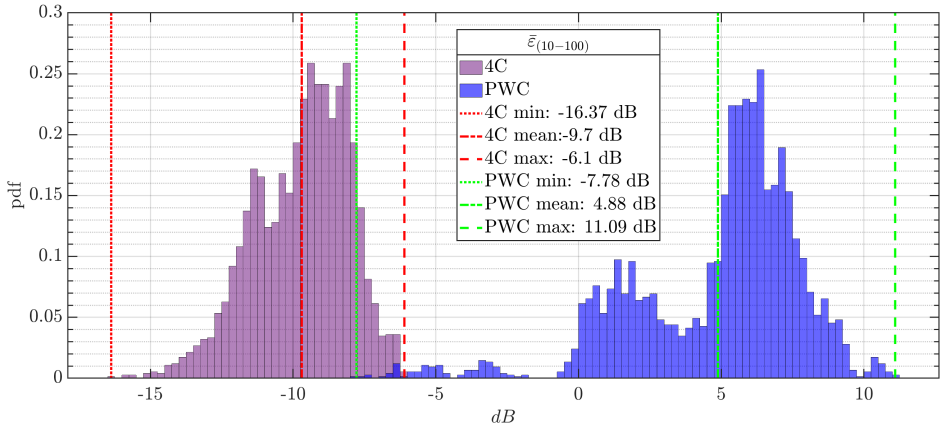


FIGURE 4.33: Histograms of $\bar{\varepsilon}$ for all defined source locations of a plane wave target sound field over the listening area, with the plane wave cancellation method (PWC) and the four-channel approach (4C).

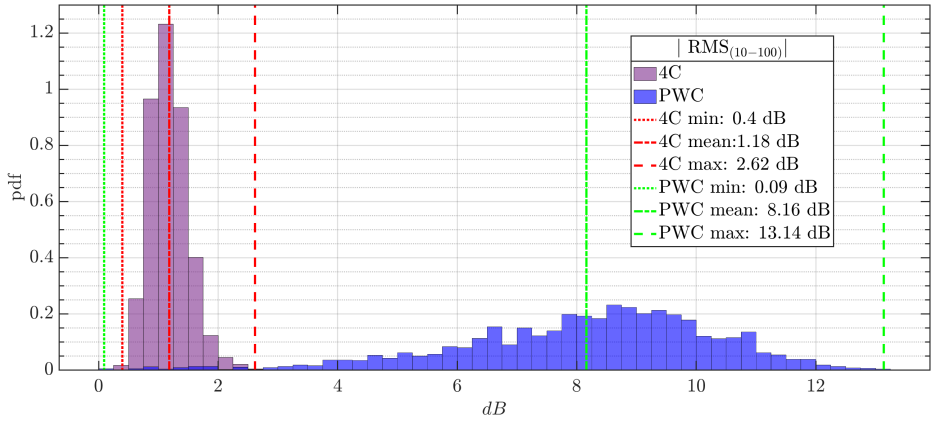


FIGURE 4.34: Histograms of $|\mathrm{RMS}_\Delta|$ for all defined source locations of the plane wave cancellation method (PWC) and the four-channel approach (4C).

band. As shown in the previous RMS data, the plane wave cancellation method outperformed the four-channel approach with a lower delta from the target of 0.09 dB versus 0.4 dB, but in the whole data set, the four-channel approach outperformed the plane wave cancellation method. The four-channel approach data had a range of $|\mathrm{RMS}_\Delta|$, 2.22 dB from the target, whereas the plane wave cancellation method had a range of 13.05 dB. The data showed that the four-channel approach worked well in all the tested locations.

Figure 4.35 shows the two histograms of the MSV from the target in the frequency band. The majority of the data for the plane wave cancellation method had an MSV value greater than the 35 dB threshold, which means that the variation between the receivers was too high for the system to have any acceptable value, as stated in Section 2.3. In contrast, with the four-channel method, the inverse is true; the majority of the data

is below the 35 dB threshold, and with a mean of 18.33 dB, most of the locations are therefore, either in the low quadrant of the acceptable range ($\leq$ 20dB), or below the 15 dB threshold and therefore denoted as good.

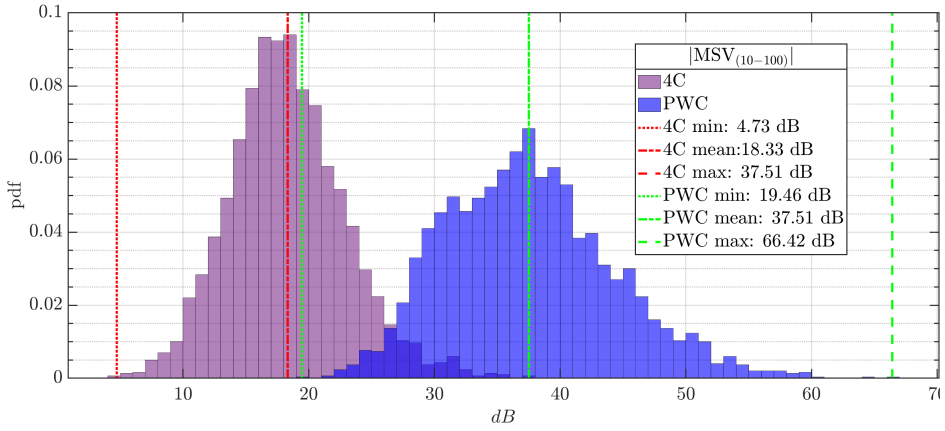


FIGURE 4.35: Histograms of MSV for all defined source locations of the plane wave cancellation method (PWC) and the four-channel approach (4C).

Appendix B gives the results of an identical study with a spherical wave target. As there was very little difference in the findings, it was decided not to include the study in the body of the thesis. The only points of note are that when the sources were in locations that gave the minimum $\bar{\varepsilon}$ value, the data showed the system was more effective at achieving a plane wave target with a value of –16.37 dB than the spherical wave target with a value of –15.67 dB. However, over all possible locations, the system was more consistent at achieving a spherical wave target, with the mean of the $\bar{\varepsilon}$ values showing –9.7 dB and –10.38 dB effectiveness at achieving a plane wave and spherical wave target, respectively. The difference between the two targets was minimal, so nothing could be gained from cross-analysis.

4.5.4 Section summary

This section reported on a number of simulation studies to understand how the control methods are affected by changes in the system arrangement and to find the limits of control. There was no point at which the four-channel approach was shown to perform, on average, worse than using the raw signal (or the plane wave cancellation method). This means the only limitation on the system is the amount of improvement possible for a given layout, room dimensions and target.

Repeating the studies from Section 3.3 to compare the plane wave cancellation method with the four-channel approach, for the cases tested here, it was shown that, in the vast majority of cases, the four-channel approach outperformed the plane wave cancellation method. Changes to the damping coefficients had little effect on the effectiveness of the

four-channel approach, and although the four-channel approach $\bar{\varepsilon}$ could be affected by both the location of the source on the wall and moving the front sources away from the wall, at worst these had a $\bar{\varepsilon}$ of –13.18 dB and at best of –19.35 dB, compared with the plane wave cancellation method values of –6.5 dB and –18.9 dB, respectively.

When targeting the plane wave sound field across the whole room, the effectiveness $\bar{\varepsilon}$ of the four-channel approach across the frequency range of 10–100 Hz dropped to –11.31 dB. Compared with the plane wave cancellation method value of –9.13 dB, the four-channel approach was clearly not limited by the whole room target.

Comparing 3,000 source location combinations using the four-channel approach method versus no control, in each layout, the $\bar{\varepsilon}$ across the frequency range of 10–100 Hz showed that no matter the layout, there was a minimum mean effectiveness of 6.1 dB. The mean of the $\bar{\varepsilon}$ values showed –9.7 dB and –10.38 dB effectiveness at achieving a plane wave and spherical wave target, respectively.

4.6 Chapter review

The chapter and review were split into four main areas.

1. **Introduction of the least-squares approaches for acoustic control and comparison with the plane wave cancellation method**
   This section introduced four methods of control to supersede the plane wave cancellation method and provided results from studies carried out using a numerical model to compare the four methods to both the plane wave cancellation method and each other. These methods were

   - **two-channel (front) least-squares approach**
   - **two-channel (rear) least-squares approach**
   - **three-channel least-squares approach**
   - **four-channel least-squares approach**

   The section showed that with both plane wave and spherical wave targets,

   - the four-channel least-squares approach performed significantly better than the plane wave cancellation method and outperformed all other methods;
   - the four-channel least-squares approach was the most effective, followed by the three-channel least-squares approach and then the two-channel (rear) least-squares approach; and
   - the two-channel (front) approach was not only the worst performer of the least-squares approaches, but it was also worse than the plane wave cancellation method.

It can confidently be said that

- when a room is set up to be ideal for the plane wave cancellation method, it is always more effective to use the four-channel, three-channel or two-channel (rear) least-squares approach than the plane wave cancellation method.

2. **Experimental studies**
   The process, methods and results from real-world studies based on the four-channel approach were detailed. It was shown that the effectiveness transfers from the analytical model to the real world and that by running the model with real-world measurement data, it can give an accurate prediction of what the final real-world result will be.

3. **Control for different room layouts**
   This section showed further results from running the analytical model for multiple layouts, which resulted in the following conclusions:

   - The four-channel, three-channel and two-channel (rear) approaches are always, on average, more effective than the plane wave cancellation method. Although the change in layout changes the effectiveness of the methods, the decline in effectiveness is deemed small enough not to be a limitation of the least-squares approach.
   - When l_1 is not the longest dimension in the environment, all the methods are less effective, but the drop in the effectiveness of the four-channel approach is less than the drop in effectiveness of the plane wave cancellation method.
   - The effectiveness of the least-squares approaches decreases as the number of areas of control increases, whereas on the whole, the plane wave cancellation method is constant, though the least-squares approaches still outperform the plane wave cancellation method.

4. **Limitations of the four-channel approach**
   This section used an analytical model to examine the four-channel approach to find its limitations, and although the effectiveness of the approach was shown to be dependent on the source location, it never became less effective than either the plane wave cancellation method or not controlling the raw signal.

It can also be confidently stated that the four-channel least-squares approach performed significantly better than the plane wave cancellation method in all the tested situations. Therefore, any acoustic control system with a four-source setup should favour the four-channel least-squares approach rather than the plane wave cancellation method.

One assumption made in this chapter that might create limitations for the four-channel least-squares approach: The area(s) under control will always be located between two

pairs of sources. There is no guarantee this will be the case in a home listening environment. Therefore further testing is needed to understand the effect of sources layout which do not adhere to this format.

Chapter 5

Predicting & controlling the sound field within an enclosure

One of the most significant issues with any consumer product that offers room correction is the end-user effort required to set up the system. Historically, systems that offer the best level of control require a higher level of data (Bharitkar and Kyriakakis (2011); Bharitkar (2003); Pedersen (2012)). This can lead to two drawbacks:

- **Time & effort**. The end user needs to work hard to get the best results, as most systems can require some (if not all) of the following:
  1. Multiple room sweep measurements at several locations
  2. Source and listener locations
  3. Room dimensions (Budd et al. (2015))
- **User error**. Depending on the user interface supplied and the amount of work the user is willing to put into the measurements, errors are highly likely.

A poorly set up system may offer a worse experience than that achieved by not using correction. The ideal system does everything possible with minimal user interaction.

In Chapter 4, it was shown that a high level of control over the listening environment can be achieved with the four-channel least-squares approach. However, this still requires several acoustic measurements – 20 for a four-source five-receiver system. This chapter focuses on understanding and predicting the sound field in a cuboid listening environment to minimise the end-user effort required to set up a system. The objective is to use real-world measurements taken with microphones built into the sources to create an accurate analytical model of a cuboid enclosure and thereby predict the frequency response at any location within the enclosure. The end user then only has

to measure at a single point to define the centre of the area to be controlled. These enhanced analytically modelled transfer functions need to be accurate enough to be used with the four-channel least-squares approach from Chapter 4.

As has already been defined in this thesis, the analytical model used is the Green's function solution to the Helmholtz equation. In order to calculate the transfer function between two points in a cuboid enclosure, the Green's function must include the coefficients for the dimensions of the enclosure, the locations of the two points of interest (source and receiver) and the damping of the enclosure surface. The theory is that, if all of the coefficients are correct for a specific real-world cuboid enclosure, the Green's function will provide an accurate model of said enclosure (Luan and Jacobsen (2008); Cox et al. (2004)). This work aims to estimate these coefficients from the real-world measurements. This chapter is split into five sections to test and prove the theory. These cover the following areas:

1. **Employing Green's function as a transfer function estimator**. An investigation is conducted to see how close the analytical model is to a real-world measurement when the coefficients are known.

2. **Estimating or compensating for damping coefficients**. Four methods are proposed and tested, two of which estimate the damping coefficient(s) within the Green's function itself and two that aim to compensate for unknown damping coefficients by filtering the Green's function.

3. **Estimating the dimensions of a cuboid enclosure**. An algorithmic process that estimates the dimensions of an enclosure from the frequency response data is introduced and tested.

4. **Estimating the source and receiver locations within a cuboid enclosure**. Using the previously estimated room dimensions, an algorithmic process using constrained least-squares and trigonometric calculations to estimate the locations (in Cartesian coordinates) from the frequency response data is introduced and tested.

5. **Using the enhanced analytical model to combine acoustic control with the four-channel least-squares approach**. Combining this chapter with the work from Chapter 4, to compare the difference in effectiveness when creating the filters for the four-channel least-squares approach to control a fixed listening area, when using either real-world measurements, the transfer functions produced using the enhanced analytical model or a combination of both.

5.1 Green's function as a transfer function estimator

As stated, the aim is to use the Green's function to predict the sound field at any location in a cuboid enclosure, where the parameters of the Green's function have been estimated for real-world measurements. To do this, the first step is to understand whether the Green's function solution can provide a method to predict these measurements within a known space.

To briefly review, in this thesis, the Green's function pressure model is defined as

$$\hat{\mathrm{P}}(\mathbf{r}|\mathbf{s}) = \sum_{n=0}^{\infty} \frac{\omega \rho_0 c_0^2 Q \ \psi_n(\mathbf{r})\psi_n(\mathbf{s})}{[2\zeta_n \omega_n \omega + i(\omega^2 - \omega_{\mathbf{n}}^2)]\mathrm{V}} \tag{5.1}$$

which, as previously stated, gives the radiated pressure at the point $\mathbf{r}_m$ due to a source at $\mathbf{s}_w$, where Q is the volume velocity of the source (unless otherwise stated, $Q = 1$), $\zeta_{\mathbf{n}}$ is the damping coefficient, V is the volume of the space in m^3 and $\psi_{\mathbf{n}}(\mathbf{r}_m)$ and $\psi_{\mathbf{n}}(\mathbf{s_w})$ are the values of the $\mathbf{n}^{th}$ mode calculated at the locations of the receiver and source respectively, such that for a cuboid room

$$\begin{aligned} \psi_{\mathbf{n}}(\mathbf{r}_m) &= \cos\frac{x_m n_1 \pi}{l_1} \cos\frac{y_m n_2 \pi}{l_2} \cos\frac{z_m n_3 \pi}{l_3}, \\ \psi_{\mathbf{n}}(\mathbf{s}_w) &= \cos\frac{\tilde{x}_w n_1 \pi}{l_1} \cos\frac{\tilde{y}_w n_2 \pi}{l_2} \cos\frac{\tilde{z}_w n_3 \pi}{l_3}, \end{aligned} \tag{5.2}$$

and $\mathbf{l} = (l_1, l_2, l_3)$ are the dimensions of the space in the form $(\mathrm{Length}, \mathrm{Width}, \mathrm{High})$. The receiver and source positions within the space are defined in three-dimensional Cartesian coordinates as $\mathbf{r}_m = (x_m, y_m, z_m)$ and $\mathbf{s}_w = (\tilde{x}_w, \tilde{y}_w, \tilde{x}_w)$, respectively, where $\mathbf{n}$ is the mode number in the form of $\mathbf{n} = (n_1, n_2, n_3)$.

The room used in Chapter 3 was no longer available for real-world measurements, it was necessary to move to a new room in the Bowers & Wilkins Southwater research office [SRO1]. This room is an acoustically treated listening environment measuring 4.66 m $\times$ 5.20 m $\times$ 2.59 m. The locations of the source and receiver used in the room throughout this chapter are provided in Table 5.1 and Table 5.3. As stated, a second set of receivers is used in this chapter; these were built into each source and are denoted by $\hat{\mathbf{r}}_{\hat{w}}$, where $\hat{w}$ is the same source as w, in that receiver $\hat{\mathbf{r}}_1$ was built into source $\mathbf{s}_1$. These receiver locations are provided in Table 5.2. They are the same $\tilde{x}$, $\tilde{y}$ as the source, but 0.3 m higher. Unless stated, any measurement taken with the built-in receivers does not include the data from the receiver built into the source and generating the measurement signal, such as in $\hat{\mathrm{P}}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)$ $\hat{w} \neq w$.

Throughout this chapter, the loudspeaker response introduced in Section 4.3.1 was taken into consideration when creating each model.

TABLE 5.1: Real-world testing loudspeaker source location.

Source No.	Source Location ($\tilde{x}$ m, $\tilde{y}$ m, $\tilde{z}$ m)	Source No.	Source Location ($\tilde{x}$ m, $\tilde{y}$ m, $\tilde{z}$ m)
$\mathbf{s}_1$	(1.004, 1.339, 0.25)	$\mathbf{s}_3$	(4.418, 1.503, 0.25)
$\mathbf{s}_2$	(1.005, 3.861, 0.25)	$\mathbf{s}_4$	(4.418, 3.690, 0.25)

TABLE 5.2: Real-world testing loudspeaker source built-in receiver location.

Built in Receiver No.	Built in Receiver Location ($\tilde{x}$ m, $\tilde{y}$ m, $\tilde{z}$ m)	Built in Receiver No.	Built in Receiver Location ($\tilde{x}$ m, $\tilde{y}$ m, $\tilde{z}$ m)
$\hat{\mathbf{r}}_1$	(1.004, 1.339, 0.55)	$\hat{\mathbf{r}}_3$	(4.418, 1.503, 0.55)
$\hat{\mathbf{r}}_2$	(1.005, 3.861, 0.55)	$\hat{\mathbf{r}}_4$	(4.418, 3.690, 0.55)

TABLE 5.3: Real-world testing microphone receiver location.

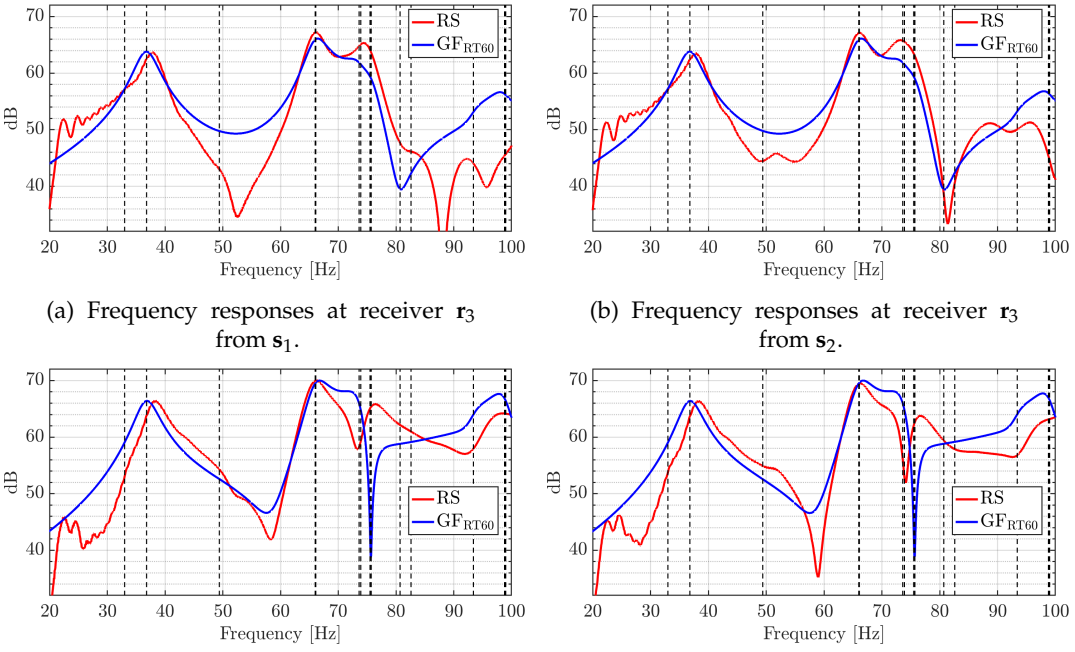
Receiver No.	Receiver Location (x m, y m, z m)	Receiver No.	Receiver Location (x m, y m, z m)
$\mathbf{r}_1$	(2.505, 1.800, 0.958)	$\mathbf{r}_4$	(3.500, 1.803, 0.958)
$\mathbf{r}_2$	(2.501, 3.392, 0.958)	$\mathbf{r}_5$	(3.496, 3.395, 0.958)
$\mathbf{r}_3$	(3.004, 2.598, 0.958)		

In this section, all the parameters are known; however, an issue arose when defining a known damping coefficient. In Section 2.1, when discussing reverberation, the Sabine reverberation time was introduced. When designing and building a listening room, a target reverberation time can be defined and approximately achieved by selecting building materials for their absorption coefficient (Everest (1989)). Furthermore, a pre-existing room can have its reverberation time changed by adding to or removing materials from surfaces or objects in the space, if both the absorption coefficient and the area of a surface are known and the damping can be calculated (Morse and Ingard (1968)). However, an accurate method of measuring the damping coefficient that an end user can use in the real world is unrealistic, meaning the damping coefficient must be calculated from an acoustic measurement. In Section 2.1, it was shown how the damping coefficient can be calculated from the RT_{60}, as used in models in Chapters 3 and 4. As presented in this chapter, the RT_{60} was calculated from the real-world measurement taken by using the receivers built into the sources. The impulse response of each measurement was used to calculate an RT_{60} value; these were then averaged to calculate a single damping coefficient using Equation 2.25.

The main issue foreseen in the RT_{60} method is that the frequency band of interest is below 100 Hz, and at low frequencies, the sound pressure does not always decay exponentially. If the decay curve deviates from a straight line on a logarithmic plot (Bruel (1951); Kuttruff (2017)), the RT_{60} may not be the best estimator.

5.1.1 Results

This section provides the results of comparing the real-world measurements RS— taken in SRO1 and the analytical model GF_{RT60}— at the receiver location $\mathbf{r}_3$, using the RT_{60} estimate from 12 real-world measurements to calculate $\zeta_{\mathbf{n}}$. The real-world transfer function measurements data from the receiver location $\mathbf{r}_3$ were not used in the calculation of $\zeta_{\mathbf{n}}$; only the data from the receiver locations $\hat{\mathbf{r}}_1$ $\hat{\mathbf{r}}_2$ $\hat{\mathbf{r}}_3$ $\hat{\mathbf{r}}_4$ were used in the analytical model.



(a) Frequency responses at receiver $\mathbf{r}_3$ from $\mathbf{s}_1$.

(b) Frequency responses at receiver $\mathbf{r}_3$ from $\mathbf{s}_2$.

(c) Frequency responses at receiver $\mathbf{r}_3$ from $\mathbf{s}_3$.

(d) Frequency responses at receiver $\mathbf{r}_3$ from $\mathbf{s}_4$.

FIGURE 5.1: Comparison of frequency responses at receiver $\mathbf{r}_3$ from each source with the real-world measurements (—RS) and the analytical model (—GF_{RT60}).

Figure 5.1 compares the frequency responses from each source at receiver $\mathbf{r}_3$ with the real-world measurements and the analytical model. The four plots in Figure 5.1 show that, when given the dimensions of an enclosure, the locations of the sources, the location of the receiver and the damping coefficients calculated from real-world RT_{60} measurements, the analytical model can provide an estimation of real-world measurements in only the broadest terms. Rough trends and significant peaks can be seen in both plots; however, none of these matches are identical to the real-world measurements. Figure 5.2 shows the average frequency response plots from both the real-world measurements and the analytical model, with the RT_{60} estimation methods at the receiver location $\mathbf{r}_3$ using all sources. The average frequency response (FA) is defined as

$$\mathrm{FA}(f) = 20\log_{10}[|\frac{1}{W}\sum_{w=1}^{W}(\tilde{p}_w(f))|], \tag{5.3}$$

where $\tilde{p}_w(f)$ is either the modelled or measured pressures at receiver points $\mathbf{r}_3(f)$.

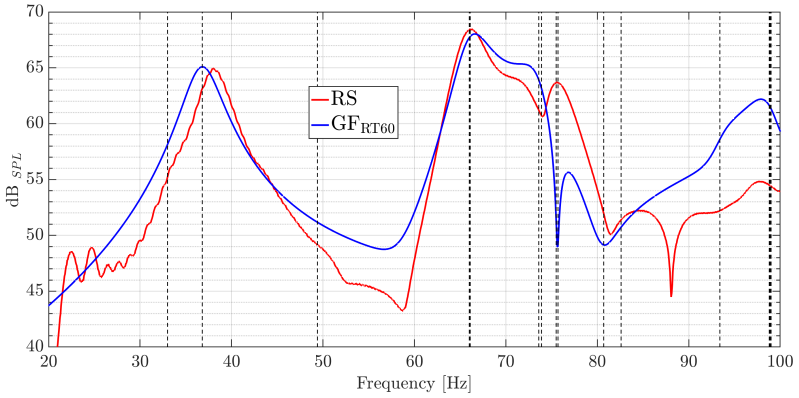


FIGURE 5.2: Comparison of FA at receiver $\mathbf{r}_3$ from each source with the real-world measurements (—RS) and the analytical model (—GF$_{RT60}$).

Figure 5.2 concurs with the data shown in Figure 5.1, confirming that again, in general terms, there is a good match between the measurements and prediction – up to approximately 75 Hz. Later in this chapter, it will be shown that the RT$_{60}$ method is not accurate enough to build effective correction filters for use with the four-channel least-squares approach. However, the RT$_{60}$ method does give a benchmark for testing against which to compare other methods. To simplify this comparison, Figure 5.3 introduces a new plot that shows the absolute delta of the FA data from the target [$|\text{FA}_\Delta|$] and the mean absolute delta of the FA from the target [$|\overline{\text{FA}}_\Delta|$] over the frequency band 20 to 100 Hz.

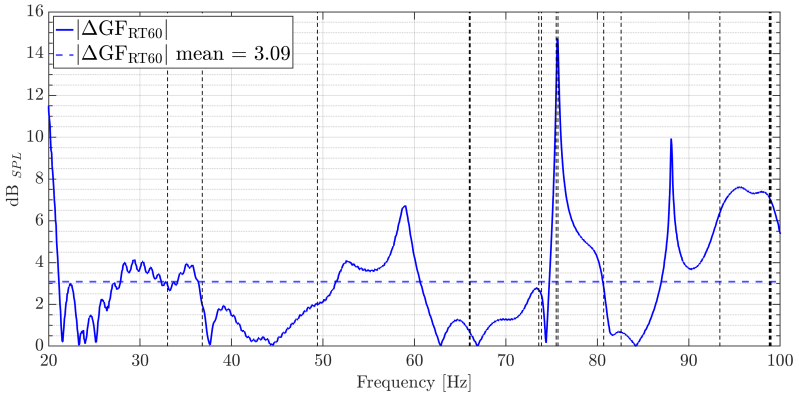


FIGURE 5.3: Comparison of $|\text{FA}_\Delta|$ at receiver $\mathbf{r}_3$ from each source with the real-world measurements (—RS) and the analytical model (—GF$_{RT60}$).

Figure 5.3 shows the $|\text{FA}_\Delta|$ for the RT$_{60}$ method with an $|\overline{\text{FA}}_\Delta|$ of 3.09 dB. To improve on the RT$_{60}$ method, new methods will need to have a lower $|\overline{\text{FA}}_\Delta|$ and show an improvement on the $|\text{FA}_\Delta|$ data.

5.2 Estimating or compensating for damping coefficients

In the last section, it was demonstrated that the analytical model can provide an estimation of real-world measurements when calculating the damping coefficients from the RT_{60}. Using this as a benchmark, four methods are introduced here to improve the estimation from the analytical model. There are two overarching approaches:

1. **Compensation**. Methods that do not estimate the damping coefficient within the Green's function but try to compensate for the unknown damping coefficient by filtering the Green's function with filters derived from real-world measurements.
2. **Estimation**. Methods that estimate the damping coefficient within the Green's function.

All the methods in this section use the real-world measurements taken using the receivers built into the sources. The methods used to estimate or compensate for the damping coefficients are as follows:

- **Compensation approaches**
    1. **Vector method**. This method groups measurement data by source and uses them to create one filter per source. The mathematics in this method are not new; however, nothing has been found that shows it being used in this way.
    2. **Matrix method**. This method takes all the measurement data to create a single $(W$ by $W)$ filter matrix that all the source inputs pass through. In theory, this should optimise the enclosure itself. Again, the mathematics in this method are not new; however, nothing has been found that shows it being used in this way.
- **Estimation approaches**
    1. **Q factor estimation method** This method uses the Q factor of the modal resonance (Everest (1989)) to estimate $\zeta_{\mathbf{n}}$.
    2. **Modal decay ($\mathbf{MT_{60}}$) method**. This method uses the MT_{60} calculation (Magalotti and Cardinali (2018)) to estimate $\zeta_{\mathbf{n}}$.

5.2.1 Compensation approaches

This section explores each of the compensation approaches in greater depth.

5.2.1.1 Vector method

Start with a single source in a cuboid environment at location $\mathbf{s}_w$. This creates a real-world impulse response, $\mathrm{P}(\mathbf{r}_m|\mathbf{s}_w)$, which is measured at a receiver elsewhere in the environment. The pressure $\mathrm{P}(\mathbf{r}_m|\mathbf{s}_w)$ can be expressed as a vector in the frequency domain, which is the transfer function between two points: $\mathbf{s}_w$ and $\mathbf{r}_m$. The Green's function given in Equation 5.25 calculates $\hat{\mathrm{P}}(\mathbf{r}_m|\mathbf{s}_w)$, modelling the transfer function between the two points $\mathbf{s}_w$ and $\mathbf{r}_m$. The objective is to create a vector in frequency v, so that

$$\mathrm{P}(\mathbf{r}_m|\mathbf{s}_w) = \hat{\mathrm{P}}(\mathbf{r}_m|\mathbf{s}_w) \cdot v. \tag{5.4}$$

For this, single-source, single-receiver example, a scalar value can be calculated for each value $v(f)$, which will give a perfect result. However, in this case, $v(f)$ is only valid for two specific points: $\mathbf{s}_w$ and $\mathbf{r}_m$. The theory behind both the vector and matrix methods is to use multiple pressure measurements to create a vector $v(f)$ (or matrix) filter that is not fixed to specific points and can be used in conjunction with the Green's function, using the locations of $\mathbf{s}_w$ and $\mathbf{r}_m$, to calculate transfer functions when there are no corresponding measurement data.

In the vector method, and using the receivers built into the source, the measurements $\mathrm{P}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)$ and Green's function $\hat{\mathrm{P}}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)$ are grouped by source, providing a single source to a multiple-receiver system. These grouped frequency responses are then split into sets according to frequency and put into vectors, so that

$$\begin{bmatrix} \mathrm{P}(\hat{\mathbf{r}}_1|\mathbf{s}_w)(f) \\ \mathrm{P}(\hat{\mathbf{r}}_2|\mathbf{s}_w)(f) \\ \vdots \\ \mathrm{P}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)(f)) \end{bmatrix} = \begin{bmatrix} \hat{\mathrm{P}}(\hat{\mathbf{r}}_1|\mathbf{s}_w)(f) \\ \hat{\mathrm{P}}(\hat{\mathbf{r}}_2|\mathbf{s}_w)(f) \\ \vdots \\ \hat{\mathrm{P}}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)(f)) \end{bmatrix} v_w(f), \tag{5.5}$$

where

$$\mathbf{p}_w(f) = \begin{bmatrix} \mathrm{P}(\hat{\mathbf{r}}_1|\mathbf{s}_w)(f) \\ \mathrm{P}(\hat{\mathbf{r}}_2|\mathbf{s}_w)(f) \\ \vdots \\ \mathrm{P}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)(f)) \end{bmatrix} \quad \& \quad \mathbf{g}_w(f) = \begin{bmatrix} \hat{\mathrm{P}}(\hat{\mathbf{r}}_1|\mathbf{s}_w)(f) \\ \hat{\mathrm{P}}(\hat{\mathbf{r}}_2|\mathbf{s}_w)(f) \\ \vdots \\ \hat{\mathrm{P}}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)(f)) \end{bmatrix}. \tag{5.6}$$

This can be solved by minimising the cost function,

$$\mathrm{J}_{vec}(f) = ||p_w(f) - \mathbf{g}_w(f)v_w(f)||^2. \tag{5.7}$$

This is an overdetermined system. Therefore, at each frequency, the least-squares solution (Elliott (2000)) is used to find v_w, so that

$$v_w = \left[\mathbf{g}_w^{\mathrm{H}}\mathbf{g}_w\right]^{-1}\mathbf{g}_w^{\mathrm{H}}\mathbf{p}_w. \tag{5.8}$$

This method only applies to fixed source locations and has to be run for each source, which means that each source has its own filter, a filter that would be non-optimal if the source were to be moved or replaced. But this method should allow a transfer function to be predicted from these fixed source locations to any receiver.

5.2.1.2 Matrix method

In the matrix method, the real-world measurements $\mathrm{P}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)$ and Green's function $\hat{\mathrm{P}}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)$ are grouped as one, providing multiple sources to a multiple-receiver system. These grouped frequency responses are then split into sets by frequency and placed into two matrices, $\mathbf{P}(f)$ and $\mathbf{G}(f)$ so that

$$\mathbf{P} = \begin{bmatrix} \mathrm{P}(\hat{\mathbf{r}}_1|\mathbf{s}_1)(f) & \mathrm{P}(\hat{\mathbf{r}}_1|\mathbf{s}_2)(f) & \dots & \mathrm{P}(\hat{\mathbf{r}}_1|\mathbf{s}_w)(f) \\ \mathrm{P}(\hat{\mathbf{r}}_2|\mathbf{s}_1)(f) & \mathrm{P}(\hat{\mathbf{r}}_2|\mathbf{s}_2)(f) & \dots & \mathrm{P}(\hat{\mathbf{r}}_2|\mathbf{s}_w)(f) \\ \mathrm{P}(\hat{\mathbf{r}}_3|\mathbf{s}_1)(f) & \mathrm{P}(\hat{\mathbf{r}}_3|\mathbf{s}_2)(f) & \dots & \mathrm{P}(\hat{\mathbf{r}}_3|\mathbf{s}_w)(f) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{P}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_1)(f)) & \mathrm{P}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_2)(f) & \dots & \mathrm{P}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)(f) \end{bmatrix} \tag{5.9}$$

and

$$\mathbf{G} = \begin{bmatrix} \hat{\mathrm{P}}(\hat{\mathbf{r}}_1|\mathbf{s}_1)(f) & \hat{\mathrm{P}}(\hat{\mathbf{r}}_1|\mathbf{s}_2)(f) & \dots & \hat{\mathrm{P}}(\hat{\mathbf{r}}_1|\mathbf{s}_w)(f) \\ \hat{\mathrm{P}}(\hat{\mathbf{r}}_2|\mathbf{s}_1)(f) & \hat{\mathrm{P}}(\hat{\mathbf{r}}_2|\mathbf{s}_2)(f) & \dots & \hat{\mathrm{P}}(\hat{\mathbf{r}}_2|\mathbf{s}_w)(f) \\ \hat{\mathrm{P}}(\hat{\mathbf{r}}_3|\mathbf{s}_1)(f) & \hat{\mathrm{P}}(\hat{\mathbf{r}}_3|\mathbf{s}_2)(f) & \dots & \hat{\mathrm{P}}(\hat{\mathbf{r}}_3|\mathbf{s}_w)(f) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathrm{P}}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_1)(f)) & \hat{\mathrm{P}}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_2)(f) & \dots & \hat{\mathrm{P}}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)(f) \end{bmatrix}, \tag{5.10}$$

This leads to the minimisation of the cost function

$$\mathrm{J}_{met}(f) = ||\mathbf{P}(f) - \mathbf{G}(f)\mathrm{Y}(f)||^2. \tag{5.11}$$

In this case, a Frobenius norm is used, as it is a matrix under evaluation; again, however, this leads to a least-squares solution being used to find $\mathrm{Y}(f)$:

$$\mathrm{Y} = \left[\mathbf{G}^{\mathrm{H}}\mathbf{G}\right]^{-1}\mathbf{G}^{\mathrm{H}}\mathbf{P}. \tag{5.12}$$

This method may be less intuitive than the vector method, as it aims to correct for the environment as a whole, with $\mathrm{Y}_{(W\times W)}$ filter the number of sources is fixed by the

method. In theory, this method can predict any number of receivers from this fixed number of sources.

5.2.2 Estimation approaches

This section explores each of the estimation approaches in greater depth.

5.2.2.1 Using the Q factor to estimate $\zeta_{\mathbf{n}}$

The Q factor ($\hat{Q}$) is a well-known method used to describe the resonance behaviour of any system with a harmonic oscillator (Elliott (2000)). In acoustics, the Q factor of individual modes in a frequency response can be defined as

$$\hat{Q} = \frac{f_c}{f_H - f_L}, \tag{5.13}$$

where f_c is the centre frequency of the mode and f_H and f_L are the frequency points above and below the centre frequency where the level drops 3 dB from the level of the centre frequency (Rumsey and McCormick (2002)). $\hat{Q}$ can then be converted to the damping coefficient (Elliott (2000))

$$\zeta_{\mathbf{n}} = \frac{1}{2\hat{Q}_{\mathbf{n}}}. \tag{5.14}$$

The first step was to find the modes within a frequency response; an example of how the modes were found is shown in Figure 5.4. First, the raw frequency response was smoothed to remove any noise; this smoothing is a simple moving average over a fixed (small) number of points within the frequency response. Double differentiation was then used to find an estimation of the peak frequencies. Although this method has not been found in any of the literature, it is just a basic implementation of standard calculus to find local maxima of a curve. The first differentiation finds the frequency of points with a gradient of zero, while the second differentiation allows any peaks to be identified. Once the peaks are estimated in the smoothed data, these are then used as a reference point to find the closest peak value in the original raw response and thereby obtain the precise modal frequencies.

Once the peaks were found, the 3 dB down point could be located for each peak and the Q factor could be calculated. $\zeta_{\mathbf{n}}$

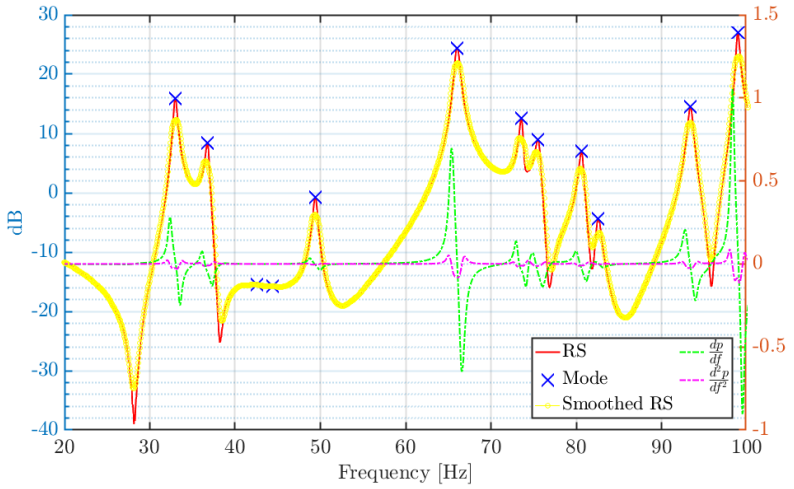


FIGURE 5.4: Double differentiation of smoothed frequency response to find modes.

5.2.2.2 $\mathbf{MT_{60}}$ estimating $\boldsymbol{\zeta_n}$

This method repurposes the works of Everest (1989), Magalotti and Cardinali (2018), Magalotti and Cardinali (2019), Magalotti (2019) and Magalotti and Ponteggia (2019) to address the low-frequency issue seen in the RT_{60} method. The MT_{60} method focuses on estimating the decay time of individual low-frequency modes, so that the damping coefficients can be calculated.

As outlined in Magalotti and Ponteggia (2019), the MT_{60} method takes in an impulse response, from which it creates a frequency response. Magalotti and Ponteggia (2019) rely on a third-party algorithm to find the modal peaks. However, this was found to be less accurate than the method outlined in Section 4.2.2.1, so the double differentiation method has been used here. Once the modal resonances have been found, an iterative process starts that aims to look at each mode in isolation. Using each mode as the centre frequency, the system applies a set of filters in the frequency domain to limit the bandwidth around the centre frequency. These bandwidths range from $\frac{1}{3}^{rd}$ octave to $\frac{1}{192}^{nd}$ octave. Then, for each bandwidth, a recursive algorithm tries to find the resonance of a single-degree-of-freedom system that gives the curve that best fits the peak. Each result of the recursive algorithm is given a score, with zero being the perfect fit. The result with the lowest score in each of the bandwidths is compared across all bandwidths to find the bandwidth-and-resonance combination with the best score. The single-degree-of-freedom system with this resonance is then transferred to the time domain, and the MT_{60} is calculated. The MT_{60} calculation is akin to RT_{60}; it is the time it takes for the single mode's linear decay to drop by 60 dB. The process outputs the scores, the MT_{60} and the dB delta of the peak-to-noise floor for each mode.

In this thesis, the impulse response for each transfer function was analysed using the above method; the results were then grouped by modal frequency. In these groups, the

data were cleaned to remove outliers by reviewing the score and dB delta peak-to-noise floor for each resonance. The remaining values were then averaged and converted to a damping coefficient using Equation 2.25.

If any mode is missing a MT_{60} value, this value is inferred from the available data. For example, if a mode missing a MT_{60} value is defined as (n_{1a},n_{2b},n_{3c}), where a, b and c denote that n_1, n_2 and n_3 are the fixed values for the mode under investigation, then (as it is assumed that the MT_{60} will change linearly with frequency) the missing value can be calculated using other available modes that satisfy either

$$\begin{aligned}
&(n_{1\hat{a}}, n_{2b}, n_{3c}) \quad \& \quad (n_{1\tilde{a}}, n_{2b}, n_{3c}) \quad \text{where} \quad n_{1a} \neq n_{1\hat{a}} \neq n_{1\tilde{a}} \\
&\text{or} \\
&(n_{1a}, n_{2\hat{b}}, n_{3c}) \quad \& \quad (n_{1\tilde{a}}, n_{2\tilde{b}}, n_{3c}) \quad \text{where} \quad n_{2b} \neq n_{2\hat{b}} \neq n_{2\tilde{b}} \\
&\text{or} \\
&(n_{1a}, n_{2b}, n_{3\hat{c}}) \quad \& \quad (n_{1a}, n_{2b}, n_{3\tilde{c}}) \quad \text{where} \quad n_{2c} \neq n_{2\hat{c}} \neq n_{2\tilde{c}}.
\end{aligned} \tag{5.15}$$

Here, $\hat{a}$, $\tilde{a}$, $\hat{b}$, $\tilde{b}$, $\hat{c}$, and $\tilde{c}$ denote their respective n_1, n_2 and n_3 values, which are fixed for the modes used as reference in each iteration of this calculation. If, for the mode under investigation, more than one of the formulae in (5.15) gives a result, these results are then averaged.

The process outlined in this section gives a damping coefficient specific to each mode. These values were then used within the Green's function as $\zeta_{\mathbf{n}}$.

5.2.3 Real-world testing results

This section shows the results from testing each method against real-world data, using the data from the RT_{60} method as a baseline. In each case, the method under test used data from the real-world measurements taken using the receivers built into the sources in room SRO1 to predict the real-world measurements at the receiver $\mathbf{r}_3$. All the other parameters for the Green's function were known, similar to the method in Section 5.1. The FA and $|FA_\Delta|$ were plotted for each test.

The RT_{60} value was used to set the damping coefficients for the vector and matrix compensation methods. As in Section 5.1 the real-world transfer function measurement data from the receiver location $\mathbf{r}_3$ were not used in the calculation of $\zeta_{\mathbf{n}}$; only the data from the receiver locations $\hat{\mathbf{r}}_1$ $\hat{\mathbf{r}}_2$ $\hat{\mathbf{r}}_3$ $\hat{\mathbf{r}}_4$ were used in the analytical model.

5.2.3.1 Vector method

This section provides the results of comparing the real-world measurements, the RT_{60} method and the vector method. The vector method is denoted as GF_v—.

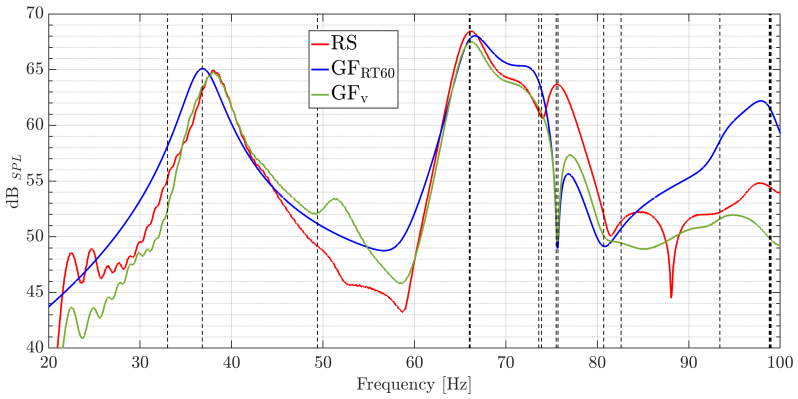


FIGURE 5.5: Comparison of FA at receiver $\mathbf{r}_3$ from each source with real-world measurements (—RS) and with two analytical models, the RT_{60} method (—GF_{RT60}) and the vector method (—GF_v).

Figure 5.5 compares the FA for the real-world measurements and the analytical models using the RT_{60} method and the vector method, with Figure 5.6 comparing the $|FA_\Delta|$ data for the analytical models using the RT_{60} method and the vector method. In Figure 5.5, it can be seen that the vector method is more effective than the RT_{60} method between 35 and 42 Hz and shows a good match to the real-world signal in the range of 65 to 75 Hz. Any differences in the two methods are easier to see in Figure 5.5.

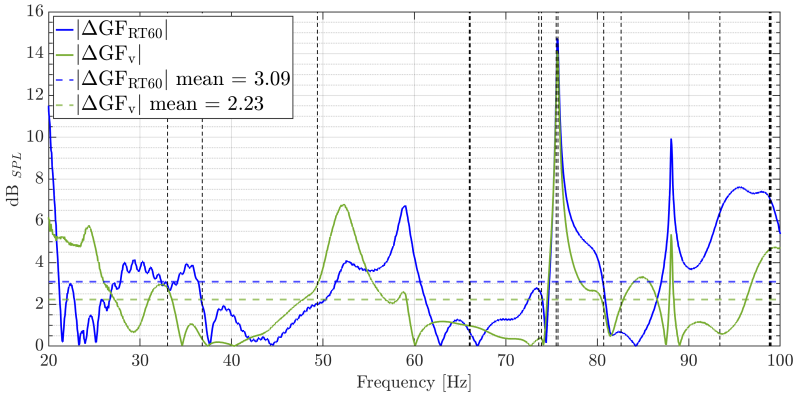


FIGURE 5.6: Comparison of $|FA_\Delta|$ at receiver $\mathbf{r}_3$, with two analytical models, the RT_{60} method (—GF_{RT60}) and the vector method (—GF_v).

In Figure 5.6, the two methods seem to offer a similar level of performance; both approximate the same peak at 75 Hz, and neither clearly outperforms the other. However, the $|\overline{FA}_\Delta|$ is 2.23 dB from the vector method, an improvement of 0.86 dB over the RT_{60}

method. A second data set analysed concerns the percentage of the method under test that is closer to the target value than the RT_{60} method. This percentage value aims to account for issues where a small set of very high or low values may bias the mean. In this case, the data show that the vector method offers an improvement over the RT_{60} method for 63.3 % of the frequency band.

5.2.3.2 Matrix method

This section shows the results of comparing the real-world measurements, the RT_{60} method and the matrix method. The matrix method is denoted as GF_m—.

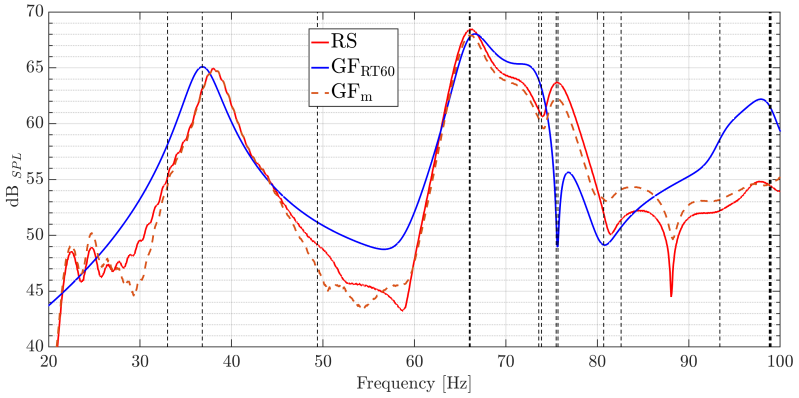


FIGURE 5.7: Comparison of FA at receiver $\mathbf{r}_3$ from each source with real-world measurements (—RS), with two analytical models, the RT_{60} method (—GF_{RT60}) and the matrix method (- - GF_m).

Figure 5.7 compares the FA for the real-world measurements and the analytical models using the RT_{60} method and the matrix method, with Figure 5.8 comparing the $|FA_\Delta|$data for the analytical models using the RT_{60} method and the matrix method. Figure 5.7 shows a clear improvement in the accuracy of the analytical model when using the matrix method, with the shape tracking of the real-world measurements being much closer than that of the RT_{60} method.

Figure 5.8 concurs with the data in Figure 5.7, with an $|\overline{FA}_\Delta|$ 1.22 dB from the matrix method, an improvement of 1.87 dB over the RT_{60} method, while the matrix method shows an improvement for 71.7 % of the frequency band. Unlike the vector method in Figure 5.6, it is clear just by looking at the data in Figure 5.8 that the matrix method is more effective than the RT_{60} method, as none of the peaks in the matrix method are comparable in size with those of the RT_{60} method, and only three of the peaks of the matrix method were worse than the $|\overline{FA}_\Delta|$ of the RT_{60} method.

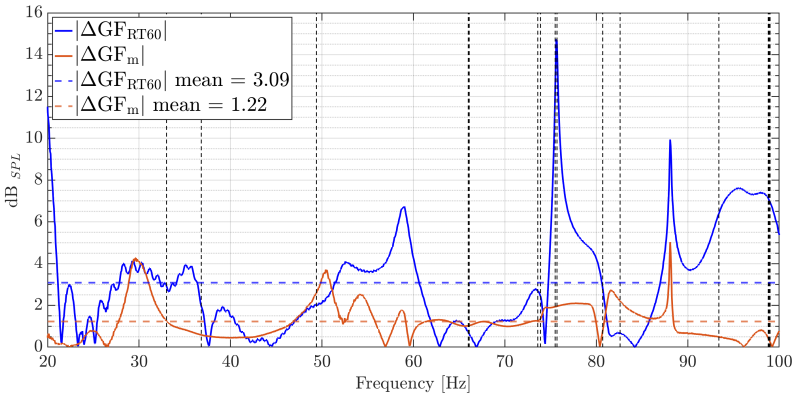


FIGURE 5.8: Comparison of $|FA_\Delta|$ at receiver $\mathbf{r}_3$, with two analytical models, the RT_{60} method (—GF_{RT60}) and the matrix method (—GF_m).

5.2.3.3 Q factor method

This section provides the results of comparing the real-world measurements, the RT_{60} method and the Q factor method. The MT_{60} method is denoted as GF_Q—.

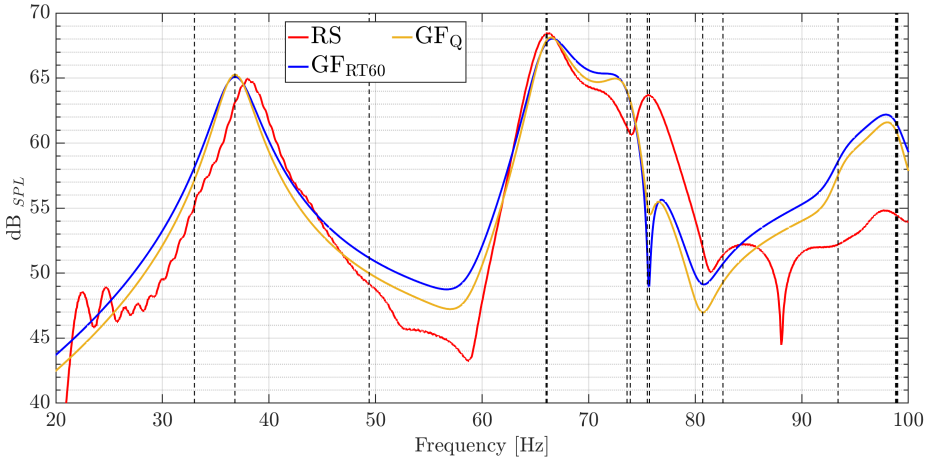


FIGURE 5.9: Comparison of FA at receiver $\mathbf{r}_3$ from each source with real-world measurements (—RS), with the two analytical models, the RT_{60} method (—GF_{RT60}) and the Q method (—GF_Q).

Figure 5.9 compares the FA for the real-world measurements and the analytical models using the RT_{60} method and the Q factor method, with Figure 5.10 comparing the $|FA_\Delta|$ for the analytical models using the RT_{60} method and the Q factor method. In Figure 5.9, a minimal difference is seen between the RT_{60} method and the Q factor method. This is shown in the data in Figure 5.10, where the Q factor method only offers an improvement over the RT_{60} method for 62.8% of the frequency band, and the $|\overline{FA}_\Delta|$ only drops by 0.41, to 2.68 dB.

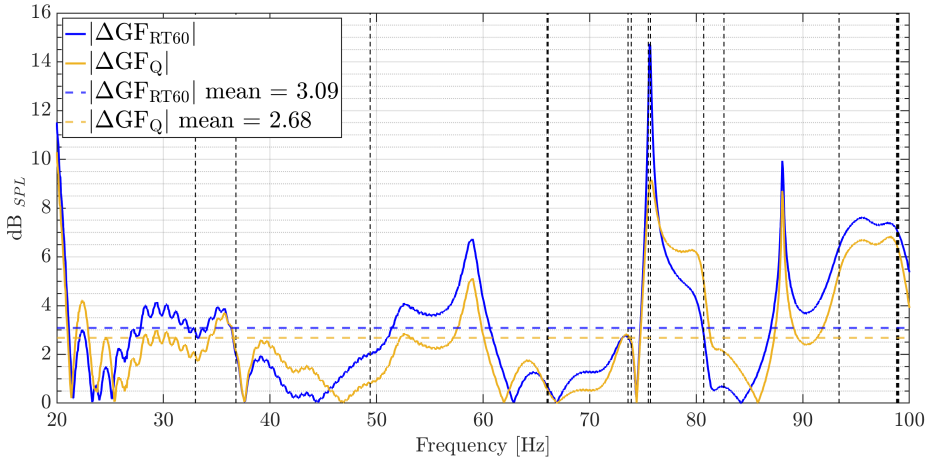


FIGURE 5.10: Comparison of $|FA_{\Delta}|$ at receiver $\mathbf{r}_3$, with two analytical models, the RT_{60} method (—GF_{RT60}) and the Q factor method (—GF_{Q}).

5.2.3.4 MT_{60} method

This section shows the results of comparing the real-world measurements, the RT_{60} method and the MT_{60} method. The MT_{60} method is denoted as GF_{MT60}—.

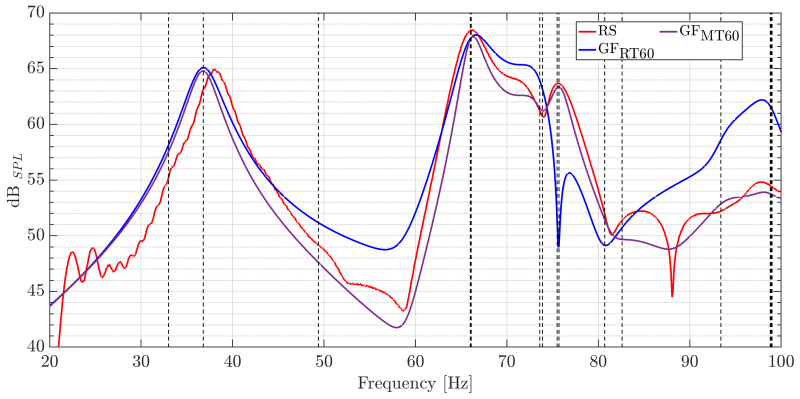


FIGURE 5.11: Comparison of FA at receiver $\mathbf{r}_3$ from each source with real-world measurements (—RS), with two analytical models, the RT_{60} method (—GF_{RT60}) and the MT_{60} method (—GF_{MT60}).

Figure 5.11 compares the FA for the real-world measurements and the analytical models using the RT_{60} method and the MT_{60} method, with Figure 5.12 comparing the $|FA_{\Delta}|$ for the analytical models using the RT_{60} method and the MT_{60} method. In Figure 5.11, it can be seen that the MT_{60} method tracks the shape of the real-world measurements much more closely than the RT_{60} method (between 38 and 82 Hz), although both methods show a similar offset on the first peak. This is also seen in Figure 5.9. It can be speculated that this offset is not due to an error in the damping estimation, but is because of an error in the measured dimension of the room or because the temperature on

the day of the measurement affected the speed of sound, meaning that 343 ms^{-1} is the wrong constant. However, multiple iterations of the model were run, adjusting both the dimensions and the speed of sound, and any situation that corrected for this offset gave a massively degraded result across the rest of the frequency band.

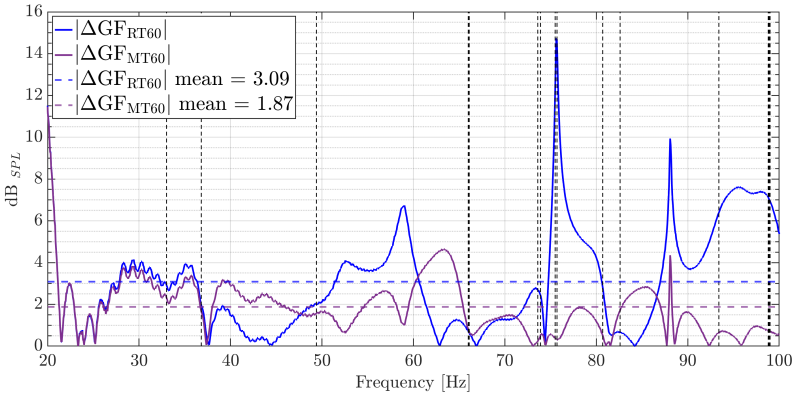


FIGURE 5.12: Comparison of $|\mathrm{FA}_\Delta|$ at receiver $\mathbf{r}_3$, with two analytical models, the RT_{60} method (—$\mathrm{GF}_{\mathrm{RT60}}$) and the MT_{60} method (—$\mathrm{GF}_{\mathrm{MT60}}$).

Figure 5.12 shows the improvement the MT_{60} method offers over the RT_{60} method. The $|\overline{\mathrm{FA}}_\Delta|$ is improved by 1.22 (to 1.87 dB) by the MT_{60} method, although the MT_{60} method is only closer to the target for 64.25 % of the frequency band. However, the MT_{60} method is closer to the target at frequencies where the RT_{60} method shows large peaks from the target.

5.2.3.5 Combined method MT_{60+}

This section combines the matrix method with MT_{60} to create a hybrid method denoted as MT_{60+}. Both methods have been shown to perform well on their own, which raises the question as to whether an improved performance could be achieved by applying the matrix method of compensation to a model after estimating the damping with MT_{60}. This is a simple method of combination, where both processes are run in series. First, the MT_{60} method tries to find the best estimation of the damping coefficients for each mode. This allows the Green's function given in (5.1) to be updated to include the estimated damping coefficients, the output of which is now denoted as $\hat{\mathrm{P}}^+(\mathbf{r}_m|\mathbf{s}_w)$. This feeds into the matrix methods in Section 5.2.1.2, such that $\mathbf{G}$ is replaced with

$$\mathbf{G}^+ = \begin{bmatrix} \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_1|\mathbf{s}_1)(f) & \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_1|\mathbf{s}_2)(f) & \dots & \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_1|\mathbf{s}_w)(f) \\ \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_2|\mathbf{s}_1)(f) & \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_2|\mathbf{s}_2)(f) & \dots & \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_2|\mathbf{s}_w)(f) \\ \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_3|\mathbf{s}_1)(f) & \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_3|\mathbf{s}_2)(f) & \dots & \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_3|\mathbf{s}_w)(f) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_1)(f)) & \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_2)(f) & \dots & \hat{\mathrm{P}}^+(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)(f) \end{bmatrix}, \tag{5.16}$$

changing the cost function to

$$J_{met^+}(f) = ||\mathbf{P}(f) - \mathbf{G}^+(f)\mathrm{Y}(f)||^2. \tag{5.17}$$

Here are the results of comparing the real-world measurements, the RT_{60} method and the MT_{60+} method. The MT_{60+} method is denoted as GF_{MT60+}—.

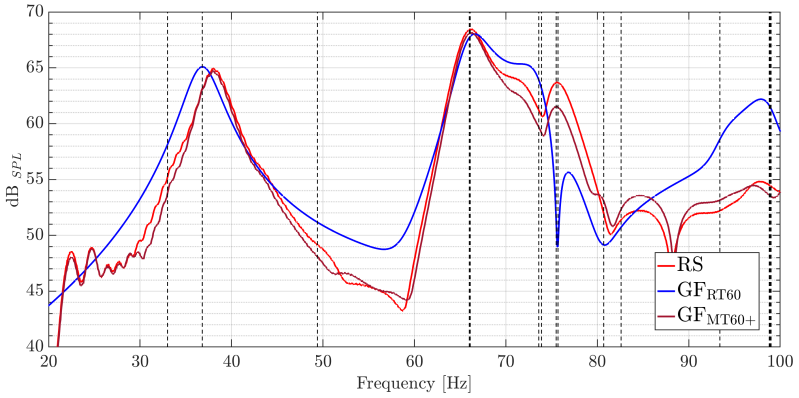


FIGURE 5.13: Comparison of FA at receiver $\mathbf{r}_3$ from each source with real-world measurements (—RS), with two analytical models, the RT_{60} method (—GF_{RT60}) and the MT_{60+} method (—GF_{MT60+}).

Figure 5.13 compares the FA for the real-world measurements and the analytical models using the RT_{60} method and the MT_{60+} method, with Figure 5.14 comparing the $|FA_\Delta|$. In Figure 5.13, it can clearly be seen that the MT_{60+} method provides a result that is very close to the real-world measurement.

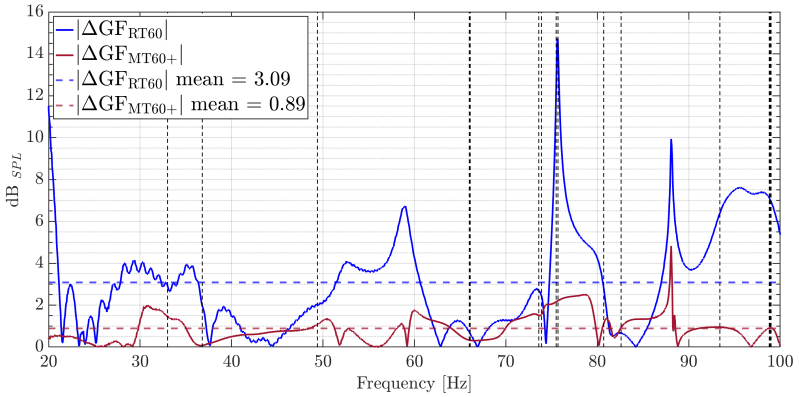


FIGURE 5.14: Comparison of $|FA_\Delta|$ at receiver $\mathbf{r}_3$ from each source with real-world measurements (—RS), with two analytical models, the RT_{60} method (—GF_{RT60}) and the MT_{60+} method (—GF_{MT60+}).

Figure 5.14 shows how close the MT_{60+} method is to the target, with the majority of the data no more than 1.5 dB from the target and an $|\overline{FA}_\Delta|$ of 0.89 dB, thus outperforming the RT_{60} method over 84.4% of the frequency band.

5.2.4 Results review

Table 5.4 brought together all the methods of estimation and compensation for the damping coefficients that were tested, showing the $|\overline{FA}_\Delta|$ for each and the percentage of the frequency band for which the method under test that is closer to the target value than the RT_{60} method. Each method outperformed the RT_{60} method.

TABLE 5.4: Comparing all methods for the estimation or compensation of damping coefficients by $|\overline{FA}_\Delta|$ and the percentage by which each method outperformed the RT_{60} method coefficient.

Method	$\lvert\overline{FA}_\Delta\rvert$ [dB]	%
GF_{RT60}	3.09	N/A
GF_{v}	2.23	63.3
GF_{m}	1.22	71.7
GF_{Q}	2.68	62.8
GF_{MT60}	1.87	64.3
GF_{MT60+}	0.89	84.4

In Table 5.4, it is clear that, of the four original methods, the matrix method performed best, with MT_{60} coming second. However, as discussed, the matrix method can be applied to the data after MT_{60}. Combining these two methods, the MT_{60+} method provided the best performance.

At this point, it was decided to move on from the damping coefficient and focus on the other variables needed for the Green's function to allow the fully enhanced analytical model to be created and tested and to determine whether the accuracy provided by the MT_{60+} method was acceptable.

5.3 Estimating the dimensions of a cuboid enclosure

To achieve the objective of understanding and predicting the sound field in a cuboid listening environment, the other parameters in the Green's function also need to be estimated, starting with the dimensions of the cuboid enclosure.

This method estimates the dimensions of a cuboid enclosure from the frequency responses. As outlined in Section 2.1, the resonance frequencies of a cuboid enclosure are

the product of the dimensions of the enclosure. Therefore, any peaks in a frequency response measured within an enclosure are created by a resonance directly linked to the dimensions of the said enclosure. The theory is that an algorithm will converge on the dimensions through a mathematical analysis of the distribution of the peaks in the measured frequency response. The higher the number of available measured frequency responses within the enclosure, the more accurate the result. There are two stages in the proposed algorithm:

Stage 1 Analysing the data by grouping peaks and calculating rough estimates for the dimensions.

Stage 2 Using estimates to create a bounded linear least-squares solution with which to re-analyse the original data for a more accurate result.

5.3.1 Stage 1

In Chapter 4, the matrix $\mathbf{C}$ was made up of individual transfer functions for each loudspeaker-to-microphone pairing. For each element in the matrix $\mathbf{C}$, the absolute values of each frequency response were calculated and analysed to find the local minima and maxima, using the same method as that used in Section 5.2.2.1. An example is shown in Figure 5.4.

The local peaks were taken from the resonance frequencies present at that location. The Q factor was also calculated for these local peaks. The Q factor is not an effective method for estimating the damping coefficient for use within the Green's function; rather, it was used here to correct the offset in the modal frequencies. The damping of a mode can change its resonance frequencies (Fahy and Walker (1998)), so each of the resonance frequencies (Elliott (2000)) was corrected using the equation

$$f_{n_{\zeta=0}} = \frac{f_{\mathbf{n}}}{\sqrt{(1 - 2\hat{Q}_{\mathbf{n}}^{-2})}}, \tag{5.18}$$

where $f_{\mathbf{n}}$ is the local peak and $f_{n_{\zeta=0}}$ is the correct value with zero damping. This process yielded an array for each element in the matrix $\mathbf{C}$. Each array contains the corrected resonance frequency values, the level of these resonance frequencies, the frequency at which the local minima values were found and the level of the local minima.

The arrays were then analysed together and used to create groups of similar frequencies. These groups contain the individual resonances found in multiple measurements. The two considerations when defining the group boundaries were as follows:

1. Any resonances must be separated by two or more local minima.

2. No group can have more resonances than the total number of original transfer functions.

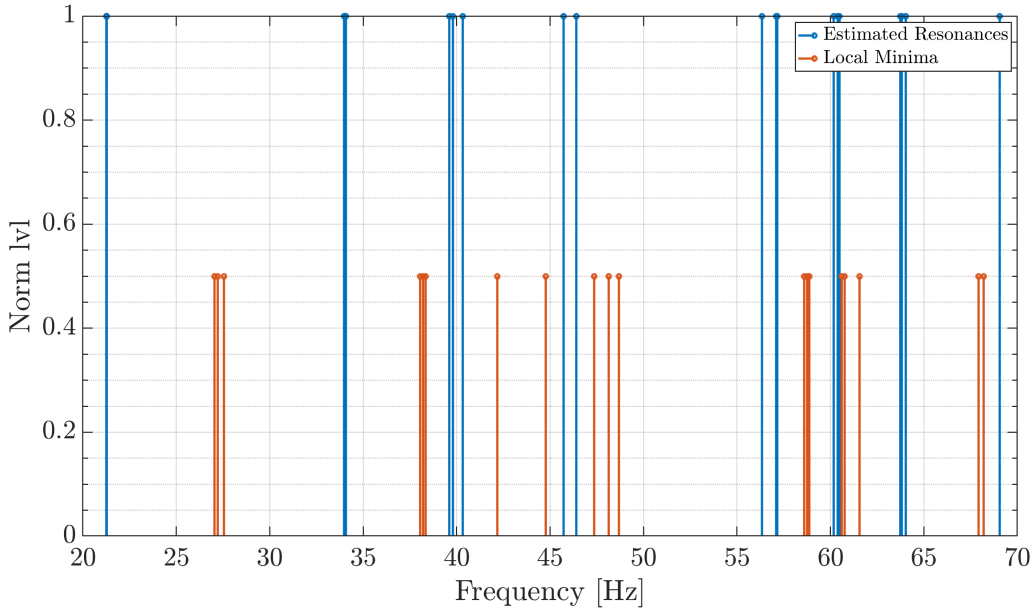


FIGURE 5.15: Example plot showing how the resonate frequencies are split into groups using the local minima.

Figure 5.15 shows a plot for the test with three transfer functions (from one loudspeaker location to three microphones). For the sake of clarity, the level of each resonance was normalised to 1.0 and that of each local minima to 0.5. This plot is a representation of how these resonances were grouped with the local minima used to define the group boundaries.

The next step was to convert these frequencies to wavelengths. The wavelengths in the first group were taken to be related to the standing wave created in the longest dimension of the space when $\mathbf{n} = (1,0,0)$. The values in this group were used to find all $\mathbf{n} = (n_1,0,0)$ groups, where $f_{\mathbf{n}} \leq f_{max}$. A cross comparison was then carried out over the individual frequencies across these groups to find the closest approximate common multiples. The system returned the common multiples from the groups with the smallest mean error. This step was repeated by taking the second group as the second-longest dimension of the space when $\mathbf{n} = (0,1,0)$. These two steps provided rough estimates of (l_1) & (l_2), which were then used to estimate the axial and tangential mode wavelengths for these two dimensions with the equation

$$\lambda_{\mathbf{n}} = \sqrt{\left(\frac{n_1}{l_1}\right)^2 + \left(\frac{n_2}{l_2}\right)^2}. \tag{5.19}$$

The first wavelength group found that was not a product of Equation 5.19 was taken to be the standing wave created in the shortest dimension of the space when $\mathbf{n} = (0, 0, 1)$, and the process to estimate its length (and error) was rerun.

The outcome of this process provided estimates of (l_1) (l_2) & (l_3) and error values for each. The smaller the error value, the higher the degree of confidence in the estimated length.

This method returned three lengths, such that $(l_1) < (l_2) < (l_3)$. It did not make any choices regarding the order of the height, length and width. In a cuboid space, the Green's function needs three-dimensional data, but the orientation is not important.

5.3.2 Stage 2

During early testing, the estimates from stage one were good, but further improvement was needed, so a second stage was introduced to the process. This stage took the estimates of (l_1) (l_2) & (l_3) and the given error values to improve the accuracy of the results.

The first step was to use the estimates (l_1) (l_2) & (l_3) and all possible combinations of $\mathbf{n}$ in which $f_{\mathbf{n}} < f_{max}$ to create a vector f_{est1_n}, where

$$f_{\text{est1}_n}(n) = \frac{c}{2}\sqrt{\left(\frac{n_1}{l_1}\right)^2 + \left(\frac{n_2}{l_2}\right)^2 + \left(\frac{n_3}{l_3}\right)^2}. \tag{5.20}$$

Then, returning to the matrix $\mathbf{C}$, the absolute values of each frequency response were summed together and analysed to find the local minima and maxima, as explained in the previous stage. The new frequency values were then placed in the vector f_{est2_n}.

The result was two vectors f_{est1_n} and f_{est2_n}, where f_{est1_n} contains all the possible modal frequencies for the enclosure calculated with the estimated values (l_1) (l_2) & (l_3), and f_{est2_n} contains all the resonance frequencies present in the summed transfer functions. By running an automated process to compare f_{est1_n} and f_{est2_n}, the corresponding values of $\mathbf{n} = (n_1, n_2, n_3)$ could be assigned to the frequencies in f_{est2_n}. Then, by rearranging the equation

$$f_{\text{est2}_n} = \frac{c}{2}\sqrt{\left(\frac{n_1}{l_{\text{est2}_1}}\right)^2 + \left(\frac{n_2}{l_{\text{est2}_2}}\right)^2 + \left(\frac{n_3}{l_{\text{est2}_3}}\right)^2} \tag{5.21}$$

into the form

$$\bar{\mathbf{f}} = \bar{\mathbf{N}}\bar{\mathbf{l}}, \tag{5.22}$$

where

$$\begin{bmatrix} \frac{2f_{\text{est2}_1}}{c}^2 \\ \frac{2f_{\text{est2}_2}}{c}^2 \\ \vdots \\ \frac{2f_{\text{est2}_N}}{c}^2 \end{bmatrix} = \begin{bmatrix} n_{1,1}^2 & n_{1,2}^2 & n_{1,3}^2 \\ n_{2,1}^2 & n_{2,2}^2 & n_{2,3}^2 \\ \vdots & \vdots & \vdots \\ n_{N,1}^2 & n_{N,1}^2 & n_{N,1}^2 \end{bmatrix} \begin{bmatrix} \frac{1}{l_{\text{est2}_1}}^2 \\ \frac{1}{l_{\text{est2}_2}}^2 \\ \frac{1}{l_{\text{est2}_3}}^2 \end{bmatrix}, \tag{5.23}$$

a bounded linear least-squares solution could be used to minimise the equation

$$||\bar{\mathbf{N}}\bar{\mathbf{l}} - \bar{\mathbf{f}}||^2. \tag{5.24}$$

The upper and lower bounds for the possible values of $\bar{\mathbf{l}}$ were set using the estimates (l_1) (l_2) & (l_3) $\pm$ as the error value for each. This extra stage either improved the accuracy of the results or returned the original estimated values. It was found to have the biggest effect on the (l_3) estimates.

5.3.3 Testing

The processes outlined in the two stages were tested with three sets of real-world data: first with the measurements from room SRO1 and then from two acoustically treated listening environments at the Bowers & Wilkins Southwater research office ([SRO2] and [SRO3]). In each case, only the data from the four sources with the four built-in receivers were used, resulting in 12 data sets. The results from these tests are shown in Table 5.5. These results show a high level of accuracy, with the largest error being 2 *cm* (rounded to the nearest *cm*).

TABLE 5.5: Comparing the measured and estimated dimensions of the real-world cuboid listening environment.

Dimension	SRO1		SRO2		SRO3	
	Measurement (m)	Estimation (m)	Measurement (m)	Estimation (m)	Measurement (m)	Estimation (m)
l_1	4.66	4.6522	7.86	7.870	8.00	7.986
l_2	5.20	5.2073	4.66	4.643	4.93	4.939
l_3	2.593	2.5750	2.59	2.596	2.96	2.947

The results in Table 5.5 furnished a high degree of confidence in this part of the process.

5.4 Estimating source and receiver locations

The previous section showed that it is possible to estimate the dimensions of a cuboid enclosure from the frequency responses. This section shows how the locations of the source and receiver can also be estimated from the frequency responses. As stated in

the last section, the distribution of the peaks in the frequency responses is related to the dimensions of the enclosure. This section works on the theory that the relative levels of the peaks (and troughs) are defined by the locations of the source and receiver, respectively, from which the transfer function is taken. Through a mathematical analysis of the levels in the transfer functions, which share either a source or receiver location, it should be possible to estimate each source and receiver location.

5.4.1 Coordinate estimation

This method is designed to estimate the coordinate locations of $\mathbf{r}_m = (x_m, y_m, z_m)$ and $\mathbf{s}_w = (\tilde{x}_w, \tilde{y}_w, \tilde{x}_w)$ by using well-known mathematical processes. No pre-existing methods have been found that use these processes in this way to estimate the coordinate locations of a source and receiver in a cuboid listening environment. This method also uses the estimated dimensions $\mathbf{l} = (l_1, l_2, l_3)$, real-world measurements, and the Green's function Equation 5.1 given as

$$\hat{\mathrm{P}}(\mathbf{r}_m|\mathbf{s}_w) = \sum_{\mathbf{n}=\mathbf{0}}^{\mathbf{N}} \frac{\omega\rho_0 c_0^2 Q \psi_{\mathbf{n}}(\mathbf{r}_m)\psi_{\mathbf{n}}(\mathbf{s}_w)}{[2\omega_{\mathbf{n}}\omega\zeta_{\mathbf{n}} + i(\omega^2 - \omega_{\mathbf{n}}^2)]\mathrm{V}}, \tag{5.25}$$

rearranging the Green's function into the form

$$\bar{\mathbf{g}} = \Theta\varphi \tag{5.26}$$

to give

$$\begin{bmatrix} \bar{g}_0 \\ \bar{g}_1 \\ \bar{g}_2 \\ \vdots \\ \bar{g}_{\dot{f}} \end{bmatrix} = \begin{bmatrix} \frac{\Omega_0\rho_0 c_0^2 Q}{[(2\omega_0\Omega_0\zeta_0 + i(\Omega_0^2 - \omega_0^2)]\mathrm{V}} & \frac{\Omega_0\rho_0 c_0^2 Q}{[(2\omega_1\Omega_0\zeta_1 + i(\Omega_0^2 - \omega_1^2)]\mathrm{V}} & \cdots & \frac{\Omega_0\rho_0 c_0^2 Q}{[(2\omega_N\Omega_0\zeta_N + i(\Omega_0^2 - \omega_N^2)]\mathrm{V}} \\ \frac{\Omega_1\rho_0 c_0^2 Q}{[(2\omega_0\Omega_1\zeta_0 + i(\Omega_1^2 - \omega_0^2)]\mathrm{V}} & \frac{\Omega_1\rho_0 c_0^2 Q}{[(2\omega_1\Omega_1\zeta_1 + i(\Omega_1^2 - \omega_1^2)]\mathrm{V}} & \cdots & \frac{\Omega_1\rho_0 c_0^2 Q}{[2\omega_N(\Omega_1\zeta_N + i(\Omega_1^2 - \omega_N^2)]\mathrm{V}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\Omega_{\dot{f}}\rho_0 c_0^2 Q}{[(2\omega_0\Omega_{\dot{f}}\zeta_0 + i(\Omega_{\dot{f}}^2 - \omega_0^2)]\mathrm{V}} & \frac{\Omega_{\dot{f}}\rho_0 c_0^2 Q}{[(2\omega_1\Omega_{\dot{f}}\zeta_1 + i(\Omega_{\dot{f}}^2 - \omega_1^2)]\mathrm{V}} & \cdots & \frac{\Omega_{\dot{f}}\rho_0 c_0^2 Q}{[(2\omega_N\Omega_{\dot{f}}\zeta_N + i(\Omega_{\dot{f}}^2 - \omega_N^2)]\mathrm{V}} \end{bmatrix} \begin{bmatrix} \psi_0(\hat{\mathbf{r}}_{\hat{w}})\psi_0(\mathbf{s_w}) \\ \psi_1(\hat{\mathbf{r}}_{\hat{w}})\psi_1(\mathbf{s}_w) \\ \vdots \\ \psi_N(\hat{\mathbf{r}}_{\hat{w}})\psi_N(\mathbf{s}_w) \end{bmatrix}. \tag{5.27}$$

Here, each row of $\bar{\mathbf{g}}$ and Θ corresponds to a given frequency shown in Θ as $\Omega_{\dot{f}}$. First, the estimated dimensions are used to calculate the modal frequencies of the enclosure, such that

$$f_{\mathbf{n}} = \frac{c_0}{2}\sqrt{\left[\frac{n_1}{l_1}\right]^2 + \left[\frac{n_2}{l_2}\right]^2 + \left[\frac{n_3}{l_3}\right]^2}, \tag{5.28}$$

where $\mathbf{n} = (n_1, n_2, n_3)$. Then, taking a set of measurement data for a point $\dot{\mathbf{p}}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)$ and creating a vector such that $\hat{\mathbf{p}} = \dot{\mathbf{p}}(\hat{\mathbf{r}}_{\hat{w}}|\mathbf{s}_w)(f_{\mathbf{n}})$, a column vector holding the pressure values at the modal frequencies was found. A sub-matrix $\hat{\Theta}$ was created from Θ. $\hat{\Theta}$

has the same columns as Θ, but only the rows where $\Omega_{\hat{f}} = 2\pi f_{\mathbf{n}}$, resulting in a matrix where each row represents a modal frequency.

In an ideal situation like $\hat{\mathbf{p}} = \hat{\Theta}\varphi$, everything in $\hat{\Theta}$ is known and φ is unknown, but with the following constraints. First, both $\hat{\mathbf{p}}$ and $\hat{\Theta}$ have to be complex values; however φ is not complex, therefore

$$\Re\{\hat{\mathbf{p}}\} = \Re\{\hat{\Theta}\}\varphi. \tag{5.29}$$

Second (ignoring phase), the magnitude of $\hat{\mathbf{p}}$ cannot be exceeded. As such,

$$|\hat{\mathbf{p}}| \geq |\hat{\Theta}|\varphi \tag{5.30}$$

and

$$\varphi(n) = \begin{cases} 1 & n = 0 \\ -1 < \varphi(n) < 1 & n > 0 \end{cases} \tag{5.31}$$

$\varphi(n)$ can be calculated, and each element in φ is equal to

$$\begin{aligned} \varphi_{\mathbf{n}} = \psi_{\mathbf{n}}(\mathbf{r}_m)\psi_{\mathbf{n}}(\mathbf{s}_w) = \cos\frac{x_m n_1 \pi}{l_1}\cos\frac{y_m n_2 \pi}{l_2}\cos\frac{z_m n_3 \pi}{l_3}\ldots \\ \cos\frac{\tilde{x}_w n_1 \pi}{l_1}\cos\frac{\tilde{y}_w n_2 \pi}{l_2}\cos\frac{\tilde{z}_w n_3 \pi}{l_3}, \end{aligned} \tag{5.32}$$

which can be solved using a selection of $\mathbf{n}$ values, where $\mathbf{n} = (n_1, n_2, n_3)$. Starting with the single-dimension modes, it is best to begin with the first two modes in any dimension, ideally at as low a frequency as possible (as the low-frequency modes give the most accurate results). For example, $\mathbf{n}_1 = (0, 1, 0)$ and $\mathbf{n}_4 = (0, 2, 0)$, meaning that both n_1 and n_3 are equal to 0. Equation 5.32 can be written as

$$\varphi_{\mathbf{n}} = \cos\frac{y_m n_2 \pi}{l_2}\cos\frac{\tilde{y}_w n_2 \pi}{l_2}, \tag{5.33}$$

and if y_m is used as the starting point, $\tilde{y}_w$ can be expressed as $\tilde{y}_w = y_m + y_{m_0}$, where both y_m and $y_m + y_{m_0}$ are $\leq l_2$. The expression for each $\varphi_{\mathbf{n}}$ can be written as

$$\varphi_1 = \cos\left(\frac{y_m \pi}{l_2}\right)\cos\left(\frac{(y_m + y_{m_0})\pi}{l_2}\right) \tag{5.34}$$

and

$$\varphi_4 = \cos\left(\frac{2y_m \pi}{l_2}\right)\cos\left(\frac{2(y_m + y_{m_0})\pi}{l_2}\right) \tag{5.35}$$

using the double-angle identity (Stroud and Booth (2001))

$$\varphi_4 = \left[2\cos^2\left(\frac{y_m\pi}{l_2}\right) - 1\right]\left[2\cos^2\left(\frac{(y_m + y_{m_0})\pi}{l_2}\right) - 1\right]. \tag{5.36}$$

Let

$$\beta = \cos\left(\frac{y_m\pi}{l_2}\right) \qquad \& \qquad \gamma = \cos\left(\frac{(y_m + y_{m_0})\pi}{l_2}\right); \tag{5.37}$$

$$\therefore\ \varphi_1 = \beta\gamma \qquad \& \qquad \varphi_4 = (2\beta^2 - 1)(2\gamma^2 - 1), \tag{5.38}$$

leading to

$$\varphi_4 = (2\beta^2 - 1)(2\frac{{\varphi_1}^2}{\beta^2} - 1) \tag{5.39}$$

which provides a quadratic in β^2. Leading to

$$2\beta^4 + (\varphi_4 - 4\varphi_1^2 - 1)\beta^2 + 2\varphi_1^2 = 0, \tag{5.40}$$

where φ_1, φ_4 and l_2 are known, which can then be solved. Once the first estimation is made, the process is then repeated in the other dimensions, $(n_1, 0, 0)$ and $(0, 0, n_3)$. These can then be confirmed and improved by conducting a cross-comparison with other measurements that share either a source or receiver location and looking at the time of flight between the source and the receiver (as outlined in Bell (2011) and Section 2.4). All the locations can be estimated by repeating this process from each source-to-receiver pairing.

5.4.2 Testing

The process outlined was tested with the data sets from three real-world rooms: SRO1; SRO2; and SRO3. In each case, two sets of data were analysed:

1. The data from the four sources with the four built-in receivers, providing the locations for both the sources Table 5.6 and the built-in receivers Table 5.7.
2. The data from the four sources to the freestanding receivers, providing the locations of the freestanding receivers Table 5.8.

The results in Tables 5.6, 5.7 and 5.8 show a high degree of confidence in this process. However, the results may be too good, as they show a maximum error of $\pm$ 2.5 cm from the measured location. The issue is that the physical measurement should have been taken from the acoustic centre of the loudspeaker driver. But this is not an easily defined point (Vanderkooy (2010)), and although every effort was made to measure each

TABLE 5.6: Comparing the measured and estimated source locations in real-world cuboid listening environments SRO1, SRO2 and SRO3.

Source	SRO1		SRO2		SRO3	
	Measurement (m)	Estimation (m)	Measurement (m)	Estimation (m)	Measurement (m)	Estimation (m)
$\mathbf{s}_1$	(1.004, 1.339, 0.25)	(0.981, 1.340, 0.229)	(1.475, 1.015, 0.850)	(1.497, 1.033, 0.858)	(2.0, 1.30, 0, 85)	(2.023, 1.285, 0, 842)
$\mathbf{s}_2$	(1.005, 3.861, 0.25)	(1.022, 3.846, 0.235)	(1.471, 3.705, 0.888)	(1.469, 3.701, 0.904)	(2.00, 3.63, 0.85)	(2.003, 3.646, 0.839)
$\mathbf{s}_3$	(4.418, 1.503, 0.25)	(4.395, 1.521, 0.227)	(6.732, 1.210, 0.840)	(6.718, 1.211, 0.825)	(7.9, 1.7, 1.15)	(7.906, 1.701, 1.137)
$\mathbf{s}_4$	(4.418, 3.690, 0.25)	(4.395, 3.701, 0.242)	(6.712, 3.519, 0.843)	(6.690, 3.543, 0.852)	(7.90, 3.23, 1.15)	(7.887, 3.233, 1.161)

TABLE 5.7: Comparing the measured and estimated built-in receiver locations in real-world cuboid listening environments SRO1, SRO2 and SRO3.

Receiver	SRO1		SRO2		SRO3	
	Measurement (m)	Estimation (m)	Measurement (m)	Estimation (m)	Measurement (m)	Estimation (m)
$\hat{\mathbf{r}}_1$	(1.004, 1.339, 0.55)	(0.981, 1.340, 0.570)	(1.475, 1.015, 1.150)	(1.497, 1.033, 1.148)	(2.0, 1.3, 1.15)	(2.023, 1.285, 1.141)
$\hat{\mathbf{r}}_2$	(1.005, 3.861, 0.55)	(1.022, 3.846, 0.561)	(1.471, 3.705, 1.188)	(1.469, 3.701, 1.204)	(2.00, 3.63, 1.15)	(2.003, 3.646, 1.142)
$\hat{\mathbf{r}}_3$	(4.418, 1.503, 0.55)	(4.395, 1.521, 0.548)	(6.732, 1.210, 1.140)	(6.718, 1.211, 1.135)	(7.9, 1.7, 1.45)	(7.906, 1.701, 1.467)
$\hat{\mathbf{r}}_4$	(4.418, 3.690, 0.55)	(4.395, 3.701, 0.527)	(6.712, 3.519, 1.143)	(6.690, 3.543, 1.155)	(7.90, 3.23, 1.45)	(7.887, 3.233, 1.4691)

TABLE 5.8: Comparing the measured and estimated freestanding receiver locations in real-world cuboid listening environments SRO1, SRO2 and SRO3.

Receiver	SRO1		SRO2		SRO3	
	Measurement (m)	Estimation (m)	Measurement (m)	Estimation (m)	Measurement (m)	Estimation (m)
$\mathbf{r}_1$	(2.505, 1.800, 0.958)	(2.505, 1.800, 0.958)	(3.566, 1.505, 0.972)	(3.581, 1.488, 0.959)	(5.0, 1.60, 1.0)	(5.000, 1.587, 0.977)
$\mathbf{r}_2$	(2.501, 3.392, 0.958)	(2.501, 3.392, 0.958)	(3.494, 3.511, 0.972)	(3.506, 3.507, 0.959)	(5.0, 3.60, 1.0)	(4.977, 3.612, 0.997)
$\mathbf{r}_3$	(3.004, 2.598, 0.958)	(3.004, 2.598, 0.958)	(4.492, 2.518, 0.946)	(4.486, 2.515, 0.952)	(6.0, 2.60, 1.0)	(6.008, 2.580, 1.018)
$\mathbf{r}_4$	(3.500, 1.803, 0.958)	(3.500, 1.803, 0.958)	(5.505, 1.508, 0.956)	(5.506, 1.487, 0.952)	(7.0, 1.60, 1.0)	(6.996, 1.590, 1.018)
$\mathbf{r}_5$	(3.496, 3.395, 0.958)	(3.496, 3.395, 0.958)	(5.505, 3.505, 0.961)	(5.512, 3.516, 0.982)	(7.0, 3.60, 1.0)	(7.014, 3.581, 1.012)

source's physical location from the acoustic centre, as in Vanderkooy (2010) the acoustic centre of a Bowers & Wilkins sealed-cabinet subwoofer was calculated to be 12.0 cm directly in front of the assumed flat-piston source, and as this work was undertaken in partnership with Bowers & Wilkins, a set of four subwoofers identical to the one used in the paper were built and used for all the measurements. The error from the measured location cannot, therefore, be seen as an error resulting from the acoustic centre of a loudspeaker driver; rather, the real accuracy test is how well data calculated in this way models the real-world measurement when used in the enhanced analytical model.

5.5 Acoustic control with the four-channel least-squares approach using an enhanced analytical room model

In this section, the three processes from this chapter are combined to build enhanced and analytically modelled transfer functions that can replace real-world measurements when building the filters required by the four-channel least-squares approach to acoustic control introduced in Chapter 4, with their effectiveness tested in real-world scenarios.

The process was tested in the real-world rooms SRO1, SRO2 and SRO3. The dimensions of each room are given in Table 5.5. Each room contains four sources with built-in receivers, the locations of which are given in Tables 5.6 and 5.7, and five freestanding receivers, the locations of which are given in Table 5.8. In each case, the 12 measurements between the sources and the built-in receivers were used to create an enhanced analytical model tailored to the specific acoustic environment. As in the previous studies in Chapters 3 and 4, the aim was to control a single listening area defined by the layout of the five freestanding receivers. A multichannel sound card was used (the RME Fireface UFX II with an the RME OctaMic II) so that all the receivers could take simultaneous measurements.

In each study in this section, four different filter banks were created for use in the four-channel least-squares approach. These were two reference banks of filters that gave a worst and best case result, and two sets of filter banks under investigation that used the enhanced analytical model. The full details of each filter bank are outlined below.

Filter bank type 1: Reference filters

(a) **Best-case filters**. These were built with the real-world measurement data taken at each freestanding receiver for the best possible result with the four-channel least-squares approach, as outlined and tested in Chapter 4. This was the target for the enhanced analytical model. Throughout this section, the filters designed in this way are referred to and denoted as 4C and plotted as (—).

(b) **Worst-case filters**. These were built using a non-enhanced analytical model, as outlined in Section 5.1, with the RT_{60} used to calculate the damping and with the dimensions and locations all known. The non-enhanced analytical model was used to create the transfer functions at four out of the five freestanding receiver locations. These locations were those of receivers r_1, r_2, r_4 and r_5, which were then combined with the real-world acoustic measurements at r_3 to create the filter. If the enhanced analytical model could not improve on this non-enhanced analytical model, there would be no reason to use the enhanced

analytical model. Throughout this section, the filters designed in this way are referred to and denoted as $4C_{RT60}$ and plotted as (—).

Filter bank type 2: Enhanced analytical model filters

(a) **Four-point filters**. Here the enhanced analytical model was used to create the transfer functions at four out of the five freestanding receiver locations. These were receivers r_1, r_2, r_4 and r_5, which were then combined with the real-world acoustic measurement at r_3 to create the filter. This approach represented the end-user taking an acoustic measurement at one point to define the centre of the listening area (tested up to approximately 4 m^2). Throughout this section, the filters designed in this way are referred to and denoted as $4C_{4MT60+}$ and plotted as (—).

(b) **Five-point filters**.The enhanced analytical model was used to create the transfer functions for all five freestanding receivers and thereby build the filter. This represents a system that can track the end-user and automatically puts them in the centre of the listening area (tested up to approximately 4 m^2), without any need for the end-user to take any acoustic measurements. Throughout this section, the filters designed in this way are referred to and denoted as $4C_{5MT60+}$ and plotted as (—) or (- -).

The process used to build the filters employs the steps outlined in Section 4.3, with the following adjustments:

1. The transfer functions from each source to all possible receivers were measured.
2. Following the above measurement of the transfer functions, they were processed to create the filters for the four-channel approach, either by using the transfer functions to the freestanding receivers to create the target 4C filters (for **Filter bank type 1:**a) or by using the transfer functions to the built-in receivers to calculate the RT_{60} for the $4C_{RT60}$ filters (for **Filter bank type 1:**b) and create the enhanced analytical model.
3. The enhanced analytical model transfer functions were then calculated and used to create the $4C_{4MT60+}$ filters (for **Filter bank type 2:**a) or combined with the data from the freestanding receiver $\mathbf{r}_3$ to create the $4C_{5MT60+}$ filters (for **Filter bank type 2:**b).
4. The four banks of filters were applied individually to the source signal and a new set of measurements was taken at the five freestanding receivers, as explained below.

In each case, both plane and spherical wave targets, as defined in Section 4.2, were used.

All studies followed these steps:

Step 1 − Measure the individual sources-to-receiver transfer functions. A swept sine-wave from 10 to 400 Hz was used to measure the signal from each loudspeaker to each measurement point. These measurements were then convolved with the inverse of the swept sine to create an impulse response.

Step 2 − Create the target. As discussed, both plane- and spherical-wave targets were used, and the required target was created in the band-limited region of 10−200 Hz.

Step 3 − Calculate the filter banks. As outlined earlier in this section, four sets of different filter banks were created using the four-channel approach outlined in Equation 4.10. These were the **best-case filter** banks using the real-world measurement data taken at each freestanding receiver, and three filter banks using modelled data (as described).

Step 4 − Each bank of filters were applied to the sine sweep to create four sets of four uniquely filtered sine sweeps, where each sweep within a set of four is assigned to a single source.

Step 5 − New measurements were taken at the five freestanding receivers using each set of filtered sine sweeps to obtain real-world results. Each measurement was obtained by simultaneously playing a single set of filtered sine sweeps through their assigned sources.

5.5.1 Results

This section provides the results from the studies in the three rooms. Figure 5.16 shows both the $\bar{\varepsilon}$ and the SA for all methods over the listening area for a frequency range of 20 to 100 Hz and with a plane-wave target in the rooms SRO1, SRO2 and SRO3. Figure 5.17 shows the same data as Figure 5.16 but with a spherical-wave target. Table 5.9 gives the $\bar{\varepsilon}$, $|\mathrm{RMS}_\Delta|$ and the MSV from the target. As in Section 4.3.2, the frequency data in the tables were analysed from 30 to 100 Hz.

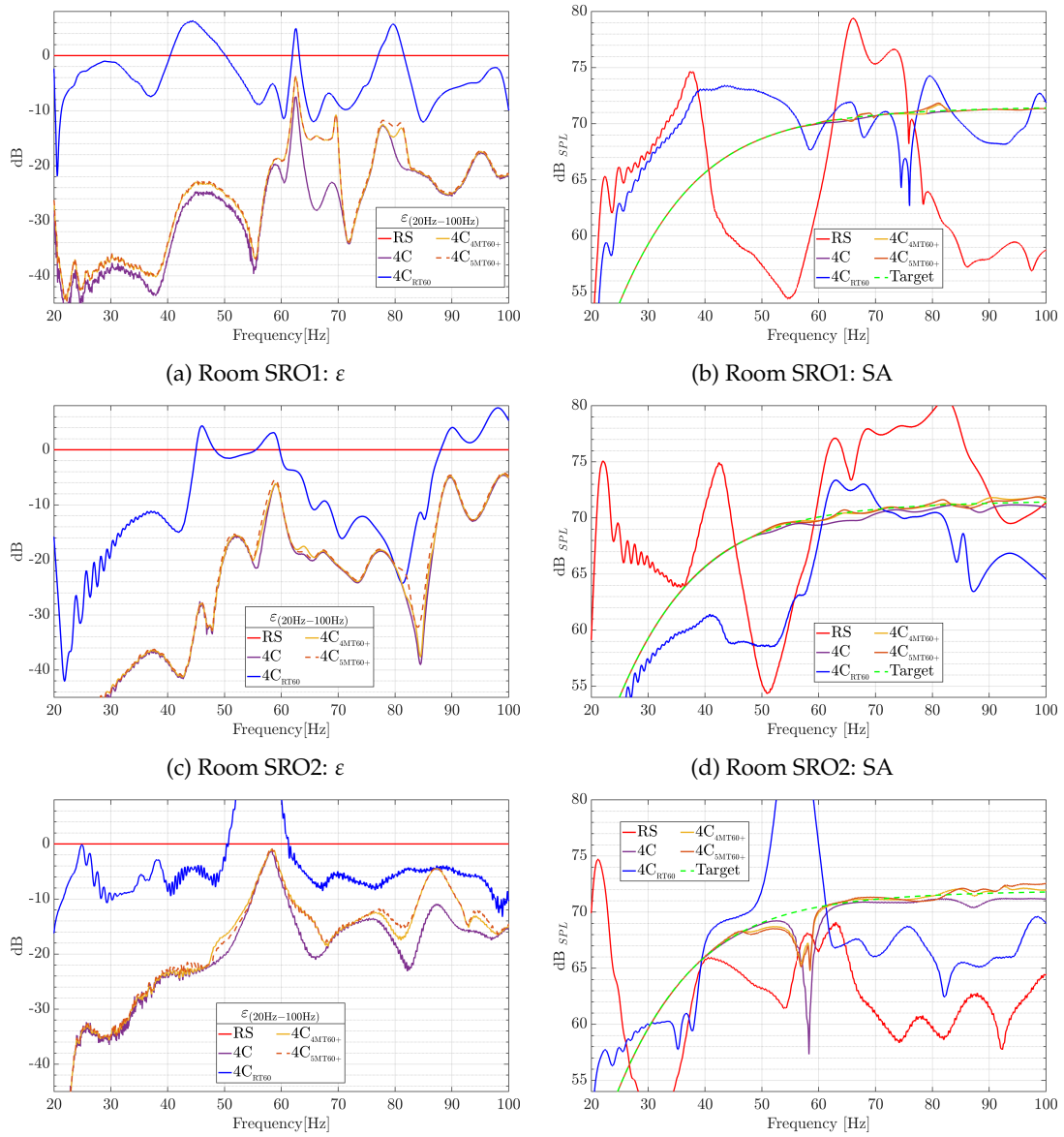
Plots 5.16a, 5.16c and 5.16e, show the $\bar{\varepsilon}$ in the rooms SRO1, SRO2 and SRO3, respectively. The first point of note is that the $4\mathrm{C}_{\mathrm{RT60}}$ filters show a poor level of effectiveness across the frequency band. This lack of control is seen in Plots 5.16b, 5.16d and 5.16f, as well as in all the plots in Figure 5.17. For all the data in Tables 5.9, the $4\mathrm{C}_{\mathrm{RT60}}$ method performed less effectively than all the other methods; in some instances, the data are even worse than the raw signal. From this, it can be seen that the non-enhanced analytical model, when the dimensions and locations are all known and the RT_{60} is used to calculate the damping, is not good enough to build a filter for use in the real world.

A comparison of the $4C_{4MT60+}$ filters with the $5C_{4MT60+}$ filters shows an almost-identical level of effectiveness in Plots 5.16a, 5.16c and 5.16e. In Plots 5.16a and 5.16e, the largest deviation between the two methods is around 80 Hz and approximately 2 dB in size. In Plot 5.16c, the largest deviation between the two methods is around 85 Hz, approximately 6 dB in size. This shows that the enhanced analytical model can provide a constant result, whether it is given the real-world acoustic measurement at r_3 or just the location of r_3. In Plots 5.16a, 5.16c and 5.16e, the $4C_{4MT60+}$ filters, the $5C_{4MT60+}$ filters and the 4C filters all show a drop in effectiveness at around 60 Hz. In each room, this coincides with the approximate wavelength of the $(0, 0, 1)$ mode in each room. Around 60 Hz, in Plots 5.16b and 5.16d, it can clearly be seen that the RS jumps from a trough to a peak. The $4C_{4MT60+}$ filters, the $5C_{4MT60+}$ filters and the 4C filters all show a good match with the target. However, Plot 5.16f is different; that is, there is a smaller trough in the 4C line, just below the approximate frequency of the $(0, 0, 1)$ mode, where the $4C_{4MT60+}$ filters, the $5C_{4MT60+}$ filters and the 4C filters do not match with the target. As in Section 4.3.2, this issue seen in Figure 5.16f will be due to the phase of the filtered signal. This raises questions as to whether, at frequencies this low, human hearing can perceive the direction of audio propagation and whether the audio would sound disjointed across the frequency band. It is commonly understood that humans have trouble locating sound sources below 200 Hz (Blauert (1996)), but this would need to be assessed through listening tests outside the remit of this thesis.

Next, the $4C_{4MT60+}$ and $4C_{5MT60+}$ filters are compared with the 4C filters. In Plot 5.16a, the $4C_{4MT60+}$ and $4C_{5MT60+}$ filters track the 4C filters almost identically, with small differences of 2 dB or less at 55 Hz to 63 Hz, and a difference of 6 dB at 85 Hz. In Plot 5.16a, the $4C_{4MT60+}$ and $4C_{5MT60+}$ filters track the 4C filters within a difference of 2 dB at 40 to 60 Hz. At 65 Hz, there is a maximum 13 dB difference, which continues up to 75 Hz; this is related to the phase issue that creates the drop in effectiveness around 60 Hz. There is about an 8 dB difference at 82 Hz. Plot 5.16e also shows that there is a large section of the frequency band at which the $4C_{4MT60+}$ and $4C_{5MT60+}$ filters are within 1 dB of the 4C, although both diverge from the 4C filters above 75 Hz.

In Plots 5.16b, 5.16d and 5.16f, it is hard, in many places, to separate the $4C_{4MT60+}$ and $4C_{5MT60+}$ filters from either 4C or the target, confirming the effectivity of the enhanced analytical model. When the $4C_{4MT60+}$, $4C_{5MT60+}$ and 4C filters do diverge, they are still within 1 dB of each other and the target in Plots 5.16b and 5.16d. In Plot 5.16f, there is the previously discussed trough just below 60 Hz; additionally, while the $4C_{4MT60+}$, $4C_{5MT60+}$ and 4C filters diverge above 80 Hz, they are still within 2 dB of each other.

The plots in Figure 5.17 show that, when the target was changed to a spherical wave, the results were very close to those seen in Figure 5.16. Again, there is very little difference in effectiveness between the $4C_{4MT60+}$ filters and $4C_{5MT60+}$ filters, nor between the $4C_{4MT60+}$ and $4C_{5MT60+}$ filters and the 4C filters.



(a) Room SRO1: ε

(b) Room SRO1: SA

(c) Room SRO2: ε

(d) Room SRO2: SA

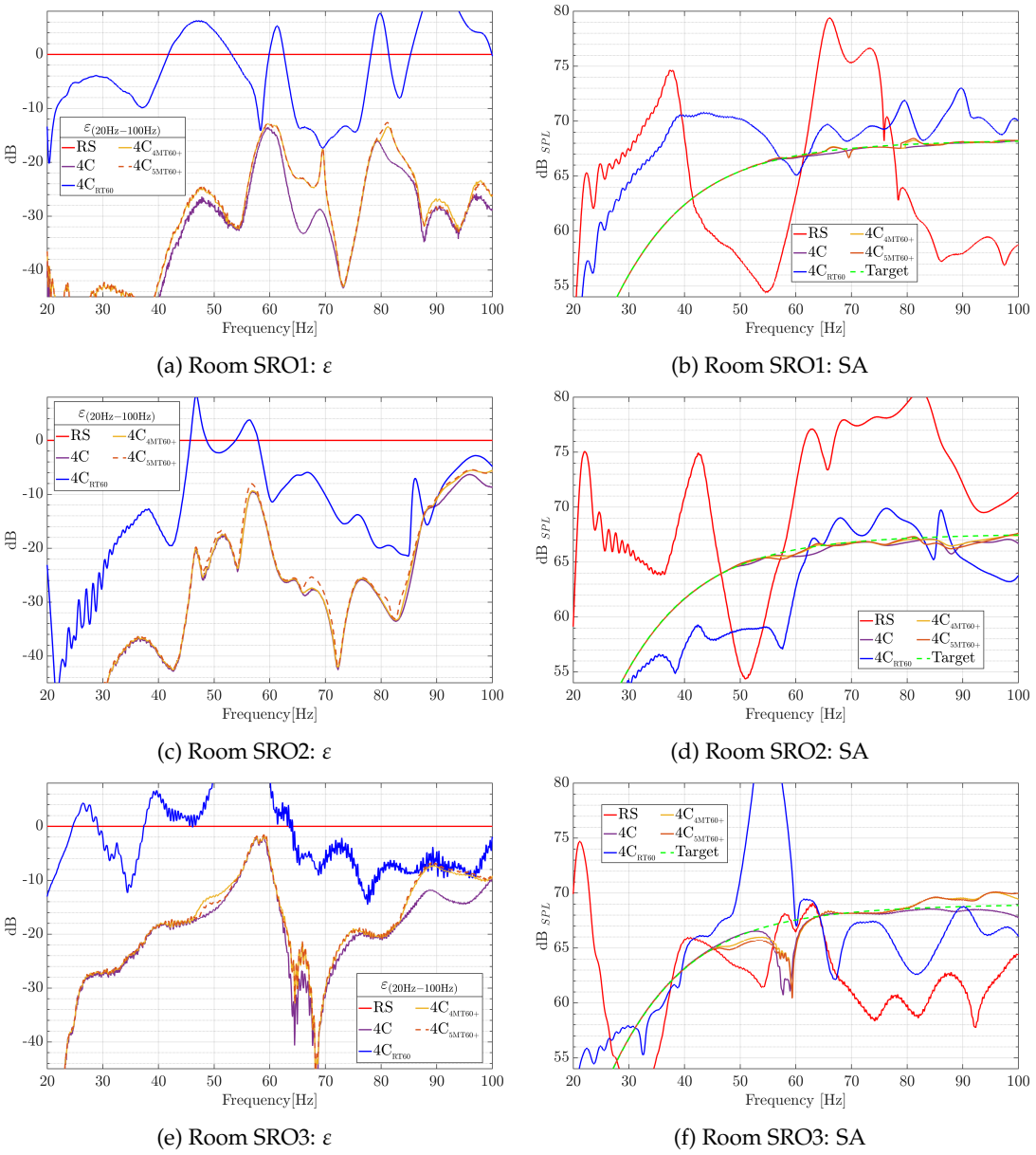
(e) Room SRO3: ε

(f) Room SRO3: SA

FIGURE 5.16: Comparing both $\bar{\varepsilon}$ and SA for all methods with a plane wave target in rooms SRO1, SRO2 and SRO3 with the raw signal (—RS), target (- -), $4C_{RT60}$ (—), 4C (—), $4C_{4MT60+}$ (—), $4C_{5MT60+}$ (- -) in the $\bar{\varepsilon}$ plots or (—) in SA plots .

In Plot 5.17a, as in Plot 5.16a, the $4C_{4MT60+}$ and $4C_{5MT60+}$ filters track the 4C filters within 2 dB above 40 Hz, except for in the 63 to 70 Hz range, where there is a maximum 10 dB difference, and in the 80 to 83 Hz range, where there is a maximum 6 dB difference. The $4C_{4MT60+}$ and $4C_{5MT60+}$ filters in Plot 5.17c are the closest to the 4C filters of any plot: no worse than a 3 dB difference at any point. Again, Plot 5.17e matches what can be seen in Plot 5.16e. There is very little difference between the $4C_{4MT60+}$, $4C_{5MT60+}$ and 4C filters in Plots 5.17b, 5.17d and 5.17f, with no more than a 3 dB difference (at worst) across all plots.

Figures 5.16 and 5.17 are confirmed by the data in Table 5.9. First, when comparing



(a) Room SRO1: ε (b) Room SRO1: SA

(c) Room SRO2: ε (d) Room SRO2: SA

(e) Room SRO3: ε (f) Room SRO3: SA

FIGURE 5.17: Comparing both $\bar{\varepsilon}$ and SA for all methods with a spherical wave target in rooms SRO1, SRO2 and SRO3 with the raw signal (—RS), target (- -), $4C_{RT60}$ (—), 4C (—), $4C_{4MT60+}$ (—), with $4C_{5MT60+}$ (- -) in the $\bar{\varepsilon}$ plots or (—) in SA plots.

$4C_{4MT60+}$ and $4C_{5MT60+}$, ε data across the full frequency shows a difference no greater than 0.45 dB in any of the rooms or with either target. The room itself seemed to determine whether the $4C_{4MT60+}$ or $4C_{5MT60+}$ filter would be more effective. In SRO1, $4C_{4MT60+}$ was more effective for both targets, whereas in SRO3, $4C_{5MT60+}$ was more effective for both targets. However, in SRO2, $4C_{4MT60+}$ was more effective with the spherical-wave target, and $4C_{5MT60+}$ was more effective with the plane-wave target. Comparing both the $4C_{4MT60+}$ and $4C_{5MT60+}$ filters with the 4C filter in ε, the data across the full frequency shows a difference of no greater than 4.24 dB in room SRO3 with the spherical-wave target. The MSV over the full frequency band also shows very little

TABLE 5.9: Frequency banded mean error [dB], RMS [dB] and MSV [dB] over the fixed listening area with both targets.

Target	Layout	Method	30 Hz $> f \geq$ 100 Hz			30 Hz $> f \geq$ 40 Hz			40 Hz $> f \geq$ 70 Hz			70 Hz $> f \geq$ 100 Hz		
			$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV
Plane Wave	SRO1	RS	0.00	2.35	36.80	0.00	11.63	40.86	0.00	2.71	51.74	0.00	0.63	31.08
		4C	-21.43	0.05	0.25	-39.99	0.00	0.06	-20.53	0.06	0.21	-19.87	0.04	0.41
		$4C_{RT60}$	-1.24	2.51	17.63	-3.80	8.28	15.21	0.48	3.13	8.99	-2.24	1.13	27.92
		$4C_{4MT60+}$	-18.62	0.11	0.65	-38.06	0.00	0.10	-16.04	0.16	1.27	-19.35	0.08	0.41
		$4C_{5MT60+}$	-18.33	0.10	0.66	-37.88	0.01	0.11	-15.92	0.15	1.28	-18.70	0.07	0.41
	SRO2	RS	0.00	6.24	22.92	0.00	7.44	74.00	0.00	5.53	20.17	0.00	6.27	9.13
		4C	-14.15	0.19	1.26	-38.67	0.01	0.03	-15.95	0.24	1.69	-11.13	0.17	1.67
		$4C_{RT60}$	-1.43	0.78	20.22	-12.83	2.63	12.18	-1.41	0.61	35.25	0.71	1.88	12.74
		$4C_{4MT60+}$	-15.06	0.18	1.10	-39.00	0.02	0.03	-15.86	0.10	1.83	-11.41	0.24	1.10
		$4C_{5MT60+}$	-14.18	0.18	1.22	-39.07	0.02	0.03	-15.01	0.12	1.74	-11.52	0.22	1.50
	SRO3	RS	0.00	13.11	38.83	0.00	11.96	43.26	0.00	12.27	28.59	0.00	17.99	40.15
		4C	-13.53	0.37	3.91	-27.71	0.01	0.09	-10.65	0.43	7.75	-14.99	0.35	2.65
		$4C_{RT60}$	13.16	12.34	30.24	-6.47	0.63	33.80	17.40	16.47	31.53	-5.96	2.90	35.54
		$4C_{4MT60+}$	-11.73	0.34	4.49	-27.75	0.01	0.09	-10.15	0.06	6.36	-10.85	0.61	5.59
		$4C_{5MT60+}$	-11.74	0.39	4.31	-26.99	0.00	0.11	-10.41	0.11	5.65	-10.60	0.71	5.81
Spherical Wave	SRO1	RS	0.00	5.17	32.59	0.00	14.45	36.65	0.00	5.53	47.53	0.00	2.19	26.88
		4C	-24.21	0.02	0.63	-47.98	0.00	0.13	-22.27	0.04	0.58	-23.77	0.02	1.01
		$4C_{RT60}$	4.03	5.07	18.46	-6.65	8.87	14.89	1.45	3.46	6.88	7.17	5.53	32.23
		$4C_{4MT60+}$	-22.21	0.10	0.81	-44.65	0.02	0.43	-20.30	0.04	0.68	-21.74	0.15	1.21
		$4C_{5MT60+}$	-21.95	0.12	0.82	-44.50	0.02	0.44	-20.30	0.04	0.69	-21.14	0.17	1.23
	SRO2	RS	0.00	9.54	16.24	0.00	10.74	67.31	0.00	8.82	13.48	0.00	9.57	2.44
		4C	-15.69	0.14	1.83	-38.88	0.00	0.45	-19.54	0.13	1.24	-12.37	0.16	3.37
		$4C_{RT60}$	-4.57	0.03	15.06	-15.69	1.66	27.43	-1.06	0.21	24.12	-8.50	0.26	6.67
		$4C_{4MT60+}$	-14.83	0.07	2.07	-38.10	0.01	0.39	-18.62	0.04	1.08	-11.14	0.10	4.21
		$4C_{5MT60+}$	-14.38	0.04	2.25	-38.30	0.01	0.38	-17.49	0.03	1.33	-10.99	0.06	4.45
	SRO3	RS	0.00	10.60	34.60	0.00	9.45	39.04	0.00	9.76	24.37	0.00	15.48	35.93
		4C	-13.35	0.34	1.77	-22.73	0.05	1.02	-11.41	0.38	5.30	-15.18	0.34	0.00
		$4C_{RT60}$	17.92	15.00	19.53	0.02	2.31	20.39	22.16	19.12	36.48	-6.77	0.20	4.80
		$4C_{4MT60+}$	-9.11	0.30	2.32	-22.57	0.09	1.03	-10.60	0.34	4.79	-6.22	0.69	1.98
		$4C_{5MT60+}$	-9.27	0.24	1.96	-22.42	0.09	1.02	-10.12	0.50	4.04	-6.26	0.68	1.76

difference between the $4C_{4MT60+}$ filters and the $4C_{5MT60+}$ filters: no more than a 0.36 dB delta, and the highest of all the MSV of 6.36 dB, well below the 15 dB, which Welti (2012) set as a benchmark for the best-performing systems.

Looking at the $|RMS_\Delta|$ over the full frequency, there are a couple of places where the $4C_{4MT60+}$ and $4C_{5MT60+}$ filters outperform the 4C filters; for example, in room SRO3 with the spherical-wave target. This was unexpected. However, as the level difference is only 0.1 dB, it would be wrong to conclude too much from this, as it may have been due to a temperature change between measurements or a small change in the source or receiver location. It is also highly unlikely that the measurements are precise enough to offer a 0.1 dB accuracy. The acoustic measurement may have given undue weight to an issue in the environment when building the 4C filters. Regularisation needs to be

improved for compensation; the enhanced analytical models automatically compensate for this, as they do not rely on local measurements in the area being controlled. In this thesis, it was more important that the $4C_{4MT60+}$ and $4C_{5MT60+}$ filters showed a comparable level of effectiveness when compared to the 4C filters.

5.6 Chapter review

This chapter has described the creation and testing of an enhanced analytical model that can replace the requirement for end-user effort in terms of taking acoustic measurements for a room compensation system. The model is based on the use of real-world data from a specific test environment to improve the accuracy with which the Green's function can predict the transfer functions for any chosen location. The chapter provides the following:

1. A study that estimates or compensates for the damping coefficients to understand the best method of accounting for specific environmentally damping information from a real-world enclosure, for use with the Green's function. It was found, when comparing the results with real-world data, that combining the MT_{60} method of estimating damping with the introduced matrix method for room compensation provided the best results.

2. The introduction and testing of a method for determining the dimensions of a cuboid room from real-world acoustic measurements taken within the room. This was shown to be effective with real-world data.

3. The introduction and testing of a method for determining the source and receiver locations within a cuboid room from real-world acoustic measurements taken within the room. This was also shown to be effective with real-world data.

This chapter has shown that the enhanced analytical model can accurately predict real-world measurements within a cuboid enclosure by analysing real-world measurements between a set of sources and receivers (where all receivers but one receiver are integrated into the sources). Furthermore, the accuracy of the predictions is such that, when used to create filters needed for the four-channel least-squares approach to room acoustic compensation, the control is of a comparable level to when real-world measurements are used to create these filters.

For the home listening environment, this means that a system needs very little or no end-user effort to achieve an effective level of acoustic control over the listening area. Once the source is in place, the system sets itself up through automated measurements between the sources and the built-in receivers. The only other needed data point is the end user's location to denote the centre of the listening zone, which has been tested up

to a size of approximately 4 m^2. This measurement can involve the end user (acoustic, manual entry, Bluetooth beacon, etc.) or not (if a visual sensor is built into the system). This system has only been shown to be effective within a cuboid enclosure and in a limited number of user cases. Any systems to market would need further study; however, this thesis establishes a strong base upon which to build.

Several assumptions have been made in this chapter which may create limitations for this work. Firstly the room under study must be cuboid. Further work would be needed to understand how well the processes outlined would transfer to a non-cuboid room. In addition, the current method for determining the dimensions searches for three distinct values, raising the question of how it would handle a situation where two or all of these dimensions were equal. Preliminary testing has taken place of a potential amendment to the algorithm which handles this case. However, further testing is needed. It is also assumed that the sources with built-in receivers are distributed around the enclosure and not co-located. The Green's function assumes a lightly damped enclosure, therefore, the enhanced analytical model assumes the same. Further work would be required to understand how the processes would work in more heavily damped spaces. However, as discussed in Chapter 2, as the damping increases in the space, the level of the resonances decrease, meaning the environment may no longer require additional acoustic control to improve the perceived audio experience.

Chapter 6

Conclusions and future work

This chapter summarises the conclusions drawn from the key findings of this thesis and discusses areas for potential further work.

6.1 Conclusions

The acoustic environment in home-listening rooms determines the audio experience perceived by the listener. The perceived audio experience relies on two main elements. The first is the environment's physical characteristics, which define the modal resonances and reverberation time, and the second is the location of the loudspeakers and listener, determining precisely how the physical characteristics will affect the audio experience. The perfect listening experience may be achieved by carefully combining these two components. However, this is unachievable for most people, as their listening environment is a multi-purpose space, meaning these components are typically compromised for ease of living, resulting in a non-ideal audio experience. This thesis has focused on low-frequency audio in cuboid rooms, investigating applying a least-squares approach to creating audio filters that can correct a non-ideal acoustic environment. This approach requires knowledge of the listening environment, which requires end-user effort in setting up the system in the real world. The research investigated methods by which the system can automatically understand the acoustic environment to build an enhanced analytical model specific to the environment under test to negate the need for this end-user effort. It then investigated if the data from the model could be used to create audio filters using the least-squares approach, which would offer a comparable improvement to a system set up manually. Although these areas are linked in the final process, the main conclusions from each area are presented individually in the following three sections of this chapter.

6.1.1 A least-squares approach to acoustic control

The study identified a least-squares approach to designing the required filters for the acoustic control as a potential solution to the issues caused by low-frequency resonances on perceived audio experience over a fixed listening area in a non-ideal listening environment. This thesis has investigated the viability of such a least-squares approach through numerical simulations and real-world experiments in a cuboid listening room. Previous research has shown that the plane-wave cancellation method offers a high level of control over resonances while also delivering a good level of subjective audio improvement. Therefore, the plane-wave cancellation method was used as a reference for the target level of resonance control.

Although highly effective, the plane-wave cancellation method is constrained by two conditions: the environment's dimensions and the placement of the loudspeakers within the environment. In many home-listening environments, it would be unrealistic to consider optimum conditions achievable. Therefore, the effect of suboptimal conditions has been investigated through numerical simulations. It has been concluded that the reduced effectiveness of the plane-wave cancellation method (under suboptimal conditions) provides adequate justification for further study into the least-squares approach.

The plane-wave cancellation method has been extensively compared with the least-squares approach for controlling room acoustics using numerical simulations. It has been concluded that when a cuboid room has the optimum conditions for adopting the plane-wave cancellation method, it is always more effective to use a four-, three- or two-channel (rear) least-squares approach when aiming to control a fixed listening area. The four-channel least-squares approach performed significantly better than the plane-wave cancellation method in all test situations, including when controlling the acoustics of the whole room. Therefore, the study concludes that any four-source setup acoustic control system should use the four-channel least-squares approach rather than the plane-wave cancellation method.

6.1.2 Automatic understanding of the acoustic environment

The Green's function has been used extensively in previous research as an analytical model for creating audio transfer functions in enclosed acoustic environments. This thesis has investigated the accuracy of the use of this model in cuboid rooms when all coefficients are known. It can be concluded that the Green's function can give an accurate approximation of real-world acoustic transfer functions in the frequency range of interest (low frequency). With this knowledge, this thesis investigated methods for extracting the coefficients required by Green's function from the analysis of real-world measurements. There are four coefficients in Green's function, based on the real-world physical properties of the modelled environments. These are: the dimensions of the

enclosure, the location of the source, the location of the receiver, and the damping coefficients of the enclosure's surfaces. Previous research has required all four of these to be measured manually and the data entered into the model. The first three physical properties can be measured in any room. The accuracy of the measurements depends on the time and effort the user is willing to take over the process and is, of course, open to user error. The fourth physical property, damping, can only be known for an enclosed space specifically designed to achieve a target level of damping. As this rarely applies in home-listening environments, this thesis focused on estimating damping.

The thesis has proposed a method for extracting the room dimensions of a cuboid environment from a frequency analysis of real-world acoustic measurements. This method groups modal resonances from multiple acoustic measurements taken within the room and shows that by mathematically analysing the frequency distribution of the resonances, the room's dimensions can be obtained with high accuracy. Therefore, it has been concluded that it is possible to establish the dimensions of a cuboid enclosed space from real-world acoustic measurements. The accuracy of the method depends on the number of acoustic measurements taken. However, in a multi-channel system with the receiver built into the sources, it has been shown that once there are four sources, the estimation can be accurate to within 2 *cm* of the physical measurements.

With the room dimensions estimation method shown to be effective, a proposed method for extracting source and receiver locations within a cuboid environment from the frequency analysis of real-world measurements has been developed. The method analyses the relative resonance levels within real-world measurements and groups these measurements by either a shared source location or a shared receiver location. The mathematical analysis of these relative levels results in a trigonometric solution from which the location can be estimated. Again, this method achieves a significant accuracy in the source and receiver locations. However, it raises the question of where the acoustic centre of the source is and how this compares to the geometrical measurement. There is an argument that estimating the locations from acoustic measurements is likely closer to the source's actual acoustic centre than when estimated from a manually geometrical measurement. The thesis's results support this. The model, which uses the previously outlined methods for estimating the source and receiver locations, shows an excellent result compared to real-world measurements. In conclusion, it has been shown that it is possible to extract the source and receiver locations within a cuboid enclosed space from real-world acoustic measurements.

Having shown it is possible to estimate the first three physical properties (the dimensions of the enclosure and the locations of the source and receiver), a more significant issue is estimating the room's damping. In a non-purpose-built environment, damping can only be estimated from loco measurements. Therefore, the thesis has compared three methods for estimating damping and introduced two compensation methods to create a modelled transfer function as close to the real-world measurement as possible.

As a reference, the RT_{60} method for damping estimation was used because this is the simplest and most straightforward way to understand an environment's acoustic profile over time. Following an extensive comparison between methods, this research has concluded that the closest results to the real-world data resulted from a combination of an estimation method, analysing of the individual modal resonances within the real-world measurements, multiple group measurements by mode number, and averages of their results to estimate damping coefficients for each mode, with a compensation method that groups all available measurement data to build a filter though the least-squares method.

6.1.3 A least-squares approach to acoustic control using modelled transfer functions

The work in the previous section on estimating the physical properties of a cuboid environment is a unique contribution to this thesis. However, the estimation of dimensions, damping, and source and receiver locations, can be used to build an enhanced analytical model that can create pseudo-real-world measurements in locations where no real-world measurements are available. This is, again, unique. These pseudo-real-world measurements are accurate enough to adequately control room acoustics in a fixed-shape listening area when used with the four-channel least-squares approach. It has been shown in three listening rooms that through either the combination of pseudo-real-world measurements with a single real-world measurement point at the centre of the listening zone or only the pseudo-real-world measurements (with a non-acoustic measurement method used to specify the centre of the listening zone), the four-channel least-squares approach can create filters that provide a level of effective control of the acoustic field. This is not only an improvement on the unfiltered sound field but is comparable to a system that has only used real-world measurement data. Therefore, a multi-channel audio system can be created that understands and predicts the sound field in a listening environment from a small set of acoustic measurement points. The system then compensates for audio quality issues at low frequencies produced in the listening area due to sub-optimum loudspeaker placements and environmental acoustic anomalies, whereby the centre of the listening area is denoted by either an acoustic measurement or known by some other method

6.2 Future work

Apart from continuing to improve the methods in Chapter 5 to investigate if the accuracy of the methods used in this study can be enhanced when using real-world data. The following are two suggested areas on which further work could focus.

6.2.1 Non-cuboid rooms

The methods used in this thesis have assumed cuboid listening rooms when estimating the physical properties of the environment and then used an enhanced analytical model to create pseudo-real-world measurements. To move into the home-listening environment, a large study must be undertaken to examine whether an enhanced analytical model can work in non-cuboid environments. Such a study could be split into two areas of investigation. First, how well does the current system work in an environment that is close to cuboid and at what point does it stop being effective? Second, can the method be transferred from a cuboid Green's function to a Green's function using arbitrary geometries?

6.2.2 Listening tests

This thesis has shown that the four-channel least-squares approach outperforms the plane-wave cancellation method when controlling modal resonances. However, the plane-wave cancellation method has been demonstrated in previous studies to deliver a good level of subjective audio improvement. This study focused on the objective metrics to show the total amount of control possible when the least-squares approach is used to design the control filters. It has previously been theorised that trying to tune the system subjectively could reduce possible control levels. Therefore, the tuning process has been left outside the remit of this thesis.

Appendix A

Multiple room layouts study

This appendix holds the information for the studies and results that are summarised in Section 4.4. Further information on the studied layouts and a detailed analysis of the data are provided. Every attempt had been made to limit repetition between this appendix and Section 4.4 to a minimum, the majority of which will be found in the description of the studies.

Several studies were carried out in order to gain a better understanding of the limitations and effectiveness of the four methods of control, these are the

- two-channel (rear) least-squares approach
- three-channel least-squares approach
- four-channel least-squares approach
- plane wave cancellation method.

Due to the amount of pre-existing work and poor performance of the two-channel (front) control approach in Section 4.2, it was removed from the testing. However, the plane wave cancellation method was still included as a benchmark.

A.1 Study setup

In this work, one room was used in two orientations, as given in Table A.1. These were called the long thin room (LT) and the short wide room (SW).

For all studies, the analytical model was the same as in Section 4.2, which used Green's function to create transfer functions where $\zeta_{\mathbf{n}} = \frac{11.052}{\omega_n}$. The filters were created as outlined in Section 4.1, with both plane and spherical wave target sound fields used.

TABLE A.1: Room orientation.

Name	l_1	l_2	l_3
Long Thin Room (LTR)	8 m	4.93 m	2.96 m
Short Wide Room (SWR)	4.93 m	8 m	2.96 m

In total, there were five layouts studied in this work, which were

1. a layout ideal for the plane wave control method, denoted as layout C
2. a layout where s_1 and s_2 are moved into a stereo placement, denoted as layout S
3. a layout where s_1 and s_2 are offset, denoted as layout O
4. a layout with dual listening areas, denoted as layout D
5. a layout with triple listening areas, denoted as layout T

A.2 Long thin room

The long thin room's dimensions are outlined in Table A.1, and for all studies, the receiver locations are given in Table 3.6. Three source configurations were studied, the locations of which are provided in Table A.2. As shown in the graphical representation in Figure A.1, they all focused on moving the front two sources, s_1 and s_2. These layouts were

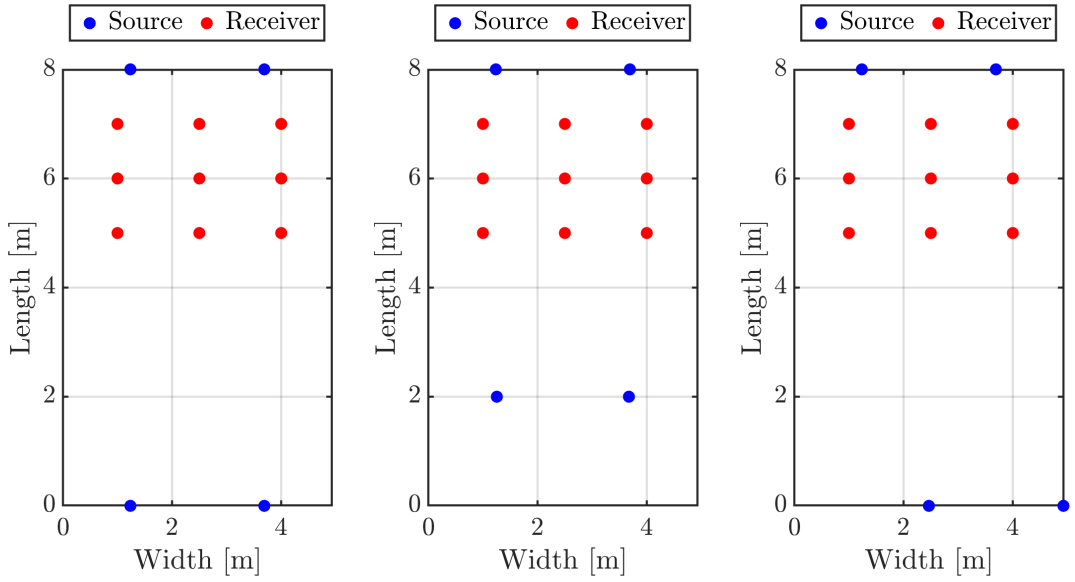
Layout 1: **LTC** (Plot A.1a) – ideal for the plane wave control method up to 69.6 Hz, as calculated using (3.7).

Layout 2: **LTS** (Plot A.1b) – s_1 and s_2 were moved into a stereo placement.

Layout 3: **LTO** (Plot A.1c) – s_1 and s_2 were offset.

TABLE A.2: Source locations of the long thin room layouts in the test.

	Source Location $(\tilde{x}$ m, $\tilde{y}$ m, $\tilde{z}$ m$)$			
	$\mathbf{s}_1$	$\mathbf{s}_2$	$\mathbf{s}_3$	$\mathbf{s}_4$
LTC	$(0.0, 1.23, 1.48)$	$(0.0, 3.70, 1.48)$	$(8.00, 1.23, 1.48)$	$(8.00, 3.70, 1.48)$
LTS	$(2.0, 1.30, 1.48)$	$(2.0, 3.63, 1.48)$	$(8.00, 1.23, 1.48)$	$(8.00, 3.70, 1.48)$
LTO	$(0.0, 2.465, 1.48)$	$(0.0, 4.93, 1.48)$	$(8.00, 1.23, 1.48)$	$(8.00, 3.70, 1.48)$



(a) LTC layout. (b) LTS layout. (c) LTO layout.

FIGURE A.1: Graphical representations of the long thin room layouts in the test.

Each of the studies in this section generated a large amount of data. Therefore, tables like A.3 compare the band-limited $\bar{\varepsilon}$, $|\mathrm{RMS}_\Delta|$ from the target, and MSV data for the set of layouts under test with both plane and spherical wave target sound fields.

To help analyse the data, a new format of graph was introduced, the first example of which can be seen in Figure A.2. These plots provided an easy way to view and compare each control method, with the layout of each set of band-limited data for the matrix under review in one plot. As with the tabled data, the same four frequency bands were given, reading from left to right. These were used to compare the relevant metrics, with the aim to keep wordy analysis to a minimum.

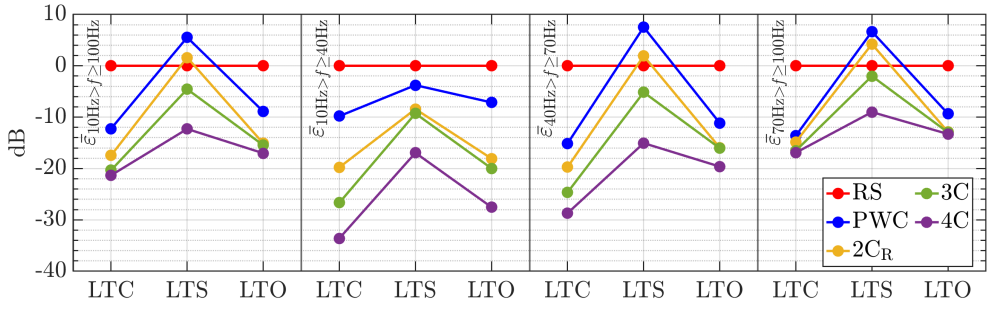


FIGURE A.2: Band-limited $\bar{\varepsilon}$ [dB] over the fixed listening area for the LTC, LTS and LTO layouts, with the plane wave target sound field; showing the raw signal (—●—RS), the plane wave cancellation method (—●—PWC), two-channel (rear) approach (- -●- -2C$_R$), three-channel approach (—●—3C) and four-channel approach (—●—4C).

TABLE A.3: Band-limited $\bar{\varepsilon}$ [dB], RMS [dB] and MSV over a fixed listening area for the LTO, LTS and LTC layouts with both plane and spherical wave target sound fields used.

Target	Layout	Method	10Hz $> f \geq$ 100Hz			10Hz $> f \geq$ 40Hz			40Hz $> f \geq$ 70Hz			70Hz $> f \geq$ 100Hz		
			$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV
Plane Wave	LTC	RS	0.00	1.75	46.17	0.00	2.14	23.93	0.00	2.67	47.72	0.00	2.72	67.00
		PWC	-12.24	0.00	5.85	-9.80	1.73	2.73	-15.13	1.21	7.85	-13.62	1.26	6.99
		$2C_R$	-17.45	0.93	1.78	-19.77	0.13	0.82	-19.74	1.36	0.84	-14.83	1.43	3.69
		3C	-20.33	0.26	1.26	-26.67	0.10	0.17	-24.65	0.28	0.44	-16.53	0.40	3.17
		4C	-21.32	0.08	1.03	-33.60	0.00	0.02	-28.69	0.02	0.25	-16.91	0.23	2.81
	LTS	RS	0.00	3.37	64.28	0.00	2.28	16.07	0.00	2.36	25.24	0.00	4.93	151.74
		PWC	5.53	6.24	23.51	-3.77	1.94	5.57	7.53	8.02	12.66	6.66	7.52	52.37
		$2C_R$	1.60	1.28	7.98	-8.42	1.64	4.13	1.92	3.12	3.57	4.23	0.26	16.29
		3C	-4.56	0.77	13.38	-9.29	0.56	4.85	-5.16	1.16	13.08	-2.01	0.62	22.25
		4C	-12.30	0.45	4.99	-16.95	0.04	0.57	-15.06	0.51	1.46	-9.01	0.84	12.96
	LTO	RS	0.00	2.32	51.71	0.00	2.79	27.76	0.00	3.02	41.34	0.00	0.83	86.21
		PWC	-8.90	1.64	12.54	-7.14	2.41	6.88	-11.16	0.94	13.28	-9.34	1.04	17.49
		$2C_R$	-15.10	0.68	4.48	-18.07	0.06	2.13	-15.94	1.05	2.48	-12.87	1.00	8.85
		3C	-15.46	0.72	4.20	-19.98	0.27	1.14	-16.05	0.93	2.46	-12.97	0.99	9.01
		4C	-17.03	0.28	3.27	-27.53	0.02	0.20	-19.62	0.18	1.52	-13.29	0.65	8.08
Spherical Wave	LTC	RS	0.00	1.75	46.17	0.00	2.14	23.93	0.00	2.67	47.72	0.00	2.72	67.00
		PWC	-8.59	0.00	4.32	-7.52	1.73	1.19	-11.22	1.21	6.32	-7.88	1.26	5.45
		$2C_R$	-12.90	1.73	1.94	-16.37	1.14	-0.24	-15.46	2.01	0.48	-9.89	2.11	5.59
		3C	-16.74	2.38	1.26	-21.94	1.70	0.01	-17.56	2.55	0.47	-14.01	2.99	3.30
		4C	-21.07	1.79	0.45	-33.52	1.70	-0.06	-22.80	1.79	-0.50	-17.50	1.89	1.92
	LTS	RS	0.00	3.37	64.28	0.00	2.28	16.07	0.00	2.36	25.24	0.00	4.93	151.74
		PWC	3.11	8.92	20.07	-2.30	0.74	2.13	5.91	10.70	9.22	2.18	10.20	48.93
		$2C_R$	-5.44	0.00	18.45	-9.82	1.58	6.09	-2.56	2.11	28.58	-7.00	2.28	20.76
		3C	-9.10	1.17	8.94	-11.13	0.30	5.38	-6.89	2.08	7.99	-10.57	1.33	13.48
		4C	-15.93	0.49	4.56	-21.24	0.09	0.70	-18.39	0.66	0.36	-12.63	0.75	12.63
	LTO	RS	0.00	2.32	50.22	0.00	2.79	26.27	0.00	3.02	39.84	0.00	0.83	84.72
		PWC	-7.75	2.35	11.04	-6.09	1.20	5.38	-9.78	2.47	11.79	-8.17	3.17	15.99
		$2C_R$	-14.04	0.71	3.41	-16.40	0.05	0.99	-15.41	1.09	0.96	-11.75	1.02	8.30
		3C	-15.39	0.18	3.55	-21.94	0.23	0.91	-16.73	0.67	1.18	-12.28	0.31	8.56
		4C	-17.59	0.34	2.75	-26.48	0.07	0.16	-19.98	0.25	0.73	-13.97	0.74	7.36

Figures A.2 and A.3 plot the band-limited $\bar{\varepsilon}$ values from Table A.3 for the plane and spherical wave target sound fields, respectively. Across all the data, it is clear that the control methods followed the expectations laid out in the hypotheses in Section 4.1.2.3, i.e. all least-squares approaches would outperform the plane wave cancellation method, and the effectiveness of the least-squares approaches would be in direct correlation to the number of degrees of freedom. All the least-squares approach plots in both figures show that the effectiveness of each method decreased as the frequencies rose and also that each method's effectiveness was more susceptible to the front source moving away

from the front wall than the source being moved along the front wall, as the LTO layout outperformed the LTS layout in each plot. It is clear that for all methods, the LTC layout was the most effective. It is interesting that with the four-channel approach, the LTS layout gave a $\bar{\varepsilon}$ of –12.30 dB, which was in line with the value from the real-world study in Table 4.5 that took place in the room these studies were modelled on.

With the LTS layout, the two-channel (rear) approach crossed the raw signal reference line when the target was a plane wave sound field, meaning the overall outcome of this approach was worse than doing nothing to the signal. Across the whole frequency, all of the least-squares approaches were more effective at achieving the spherical wave target in Figure A.3 than the spherical wave target in Figure A.2, apart from when using the LTC Layout.

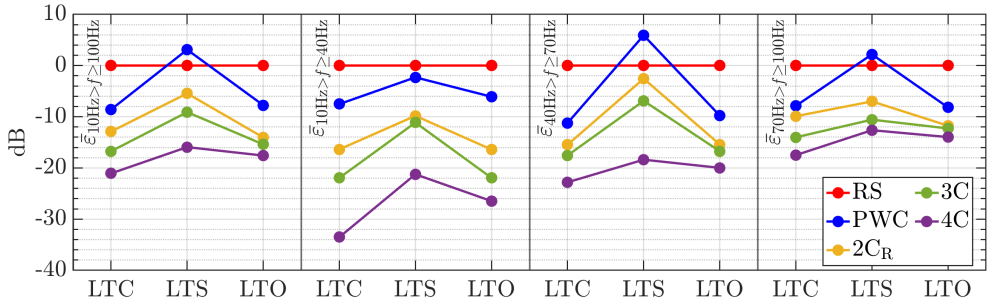


FIGURE A.3: Band-limited $\bar{\varepsilon}$ [dB] over the fixed listening area for the LTC, LTS and LTO layouts with the spherical wave target sound field; showing the raw signal (—●—RS), the plane wave cancellation method (—●—PWC), two-channel (rear) approach (- - ●- -2C$_R$), three-channel approach (—●—3C) and four-channel approach (—●—4C).

In Table A.3, the $|\text{RMS}_\Delta|$ data showed that, apart from when using the plane wave cancellation method in the LTC setup analysing the entire frequency band (where the error was zero) for the reasons already discussed in Chapter 3, the three-channel and four-channel least-squares approaches showed an improvement on the plane wave cancellation method. Furthermore, all least-squares approaches, except LTS with the two-channel (rear) plane wave and the raw signal full frequency band, showed an improvement.Table A.3 shows that the MSV data for all methods of control are below the acceptable threshold of 35 dB; also, it is clear that the four-channel approach consistently gives MSV values that were always below the 15 dB threshold, denoting a good system.

All the data in this set of studies agreed with the data already seen in Chapter 4, namely that in order of effectiveness, the four-channel, three-channel and two-channel (rear) approaches were more effective than the plane wave cancellation method. Although the change in layout did change the effectiveness of the methods, nothing has arisen so far to point to a limitation of the least-squares approach, especially when compared with the plane wave cancellation method. Due to this fact, when considering four

sources, the four-channel approach will always be the most effective least-squares option.

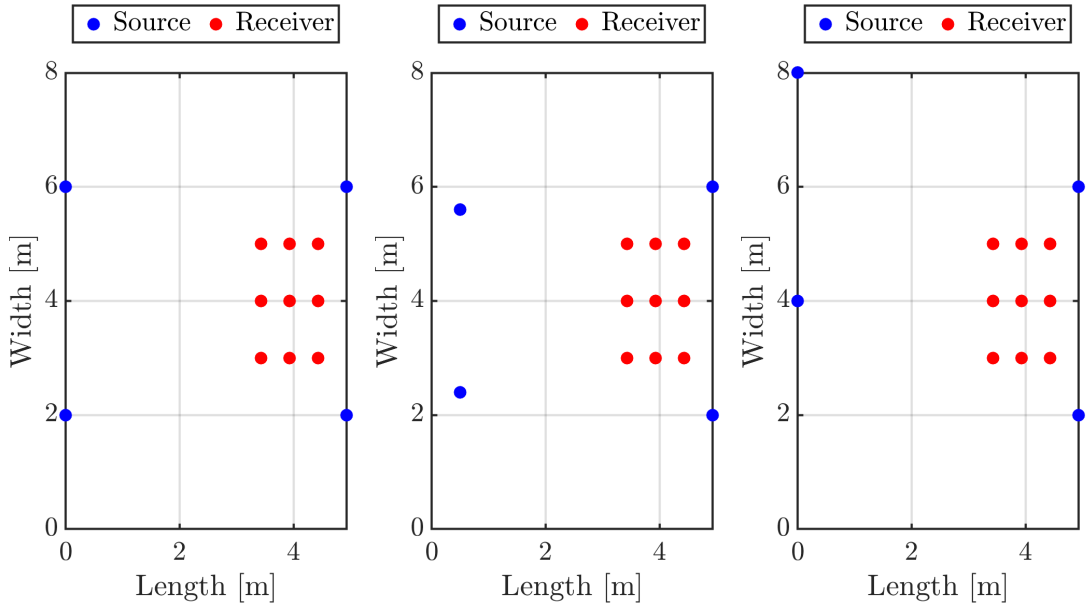
A.3 Short wide room

The reason for the short wide room was to understand how effective these systems can be when l_1 is not the largest dimension in the environment. The short wide room dimensions are outlined in Table A.1, and for all studies, the receiver locations are given in Table A.4. Again, three source configurations were studied, and the locations are given in Table A.5. These are shown in the graphical representation in Figure A.4. They all focused on moving the two front sources, s_1 and s_2. The layouts were

Layout 1: **SWC** (Plot A.4a) – ideal layout for plane wave control. However, the increased distance between the sources should lower the maximum frequency of a plane wave that can be created (as outlined in Section 3.1) to 42.9 Hz, as calculated using (3.7).

Layout 2: **SWS** (Plot A.4b) – s_1 and s_2 are moved into a stereo placement.

Layout 3: **SWO** (Plot A.4c) – s_1 and s_2 are offset.



(a) SWC. (b) SWS. (c) SWO.

FIGURE A.4: Short wide room layouts.

In this section, Table A.6 shows the band-limited $\bar{\varepsilon}$, $|\text{RMS}_\Delta|$ from the target, and MSV data for the STO, STS and STC layouts with both plane and spherical wave target sound fields for all methods in the test.

TABLE A.4: Short wide room receiver locations.

Receiver No.	Receiver Location (x m, y m, z m)	Receiver No.	Receiver Location (x m, y m, z m)	Receiver No.	Receiver Location (x m, y m, z m)
$\mathbf{r}_1$	(3.43, 3.00, 1.48)	$\mathbf{r}_4$	(3.93, 3.00, 1.48)	$\mathbf{r}_7$	(4.43, 3.00, 1.48)
$\mathbf{r}_2$	(3.43, 4.00, 1.48)	$\mathbf{r}_5$	(3.93, 4.00, 1.48)	$\mathbf{r}_8$	(4.43, 4.00, 1.48)
$\mathbf{r}_3$	(3.43, 5.00, 1.48)	$\mathbf{r}_6$	(3.93, 5.00, 1.48)	$\mathbf{r}_9$	(4.43, 5.00, 1.48)

TABLE A.5: Short wide room source locations.

Layout	Source Location ($\tilde{x}$ m, $\tilde{y}$ m, $\tilde{z}$ m)			
	$\mathbf{s}_1$	$\mathbf{s}_2$	$\mathbf{s}_3$	$\mathbf{s}_4$
SWC	(0.0, 2.0, 1.48)	(0.0, 6.0, 1.48)	(4.93, 2.0, 1.48)	(4.93, 6.0, 1.48)
SWS	(0.5, 1.25, 1.48)	(0.5, 3.68, 1.48)	(4.93, 2.0, 1.48)	(4.93, 6.0, 1.48)
SWO	(0.0, 2.4, 1.48)	(0.0, 5.6, 1.48)	(4.93, 2.0, 1.48)	(4.93, 6.0, 1.48)

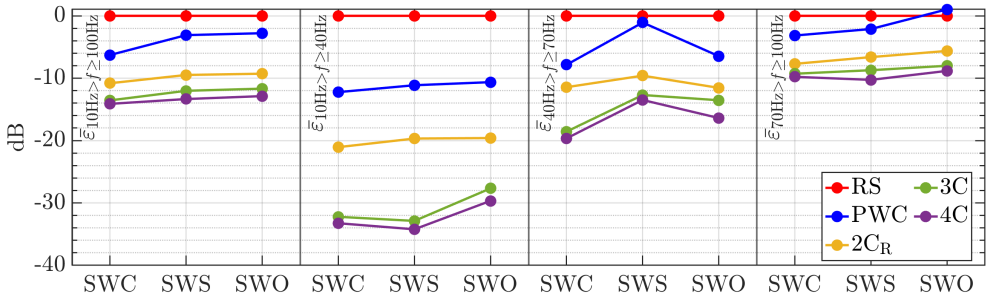


FIGURE A.5: Band-limited $\bar{\varepsilon}$[dB] over the fixed listening area for the SWC, SWS and SWO layouts, with plane wave target sound fields; with the raw signal (—●—RS), the plane wave cancellation method (—●—PWC), two-channel (rear) approach (- - ●- -2C$_R$), three-channel approach (—●—3C) and four-channel approach (—●—4C).

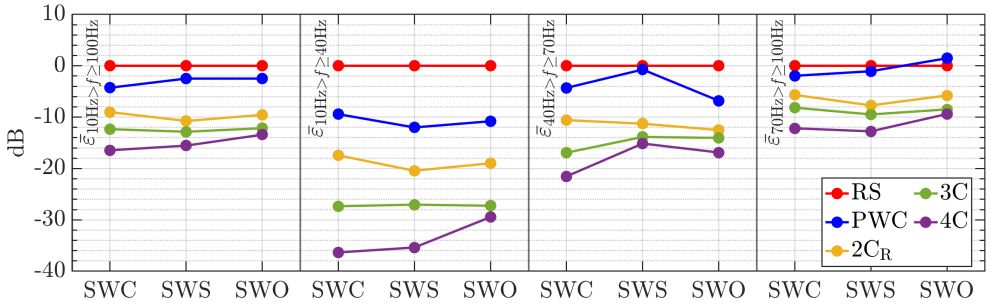


FIGURE A.6: Band-limited $\bar{\varepsilon}$ [dB] over the fixed listening area for the SWC, SWS and SWO layouts, with spherical wave target sound fields; with the raw signal (—●—RS), the plane wave cancellation method (—●—PWC), two-channel (rear) approach (- - ●- -2C$_R$), three-channel approach (—●—3C) and four-channel approach (—●—4C).

Figure A.5 and Figure A.6 plot the band-limited $\bar{\varepsilon}$ values from Table A.6 for the plane and spherical wave target sound fields, respectively. In these plots, it can be seen that the four-channel approach was more effective at achieving the spherical wave target over the full bandwidth. All methods showed a similar level of effectiveness for both

TABLE A.6: Band-limited $\bar{\varepsilon}$ [dB], RMS [dB], and MSV over a fixed listening area for the SWO, SWS and SWC layouts with both plane and spherical wave target sound fields used.

Target	Layout	Method	10Hz $> f \geq$ 100Hz			10Hz $> f \geq$ 40Hz			40Hz $> f \geq$ 70Hz			70Hz $> f \geq$ 100Hz		
			$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV
Plane Wave	SWC	RS	0.00	3.88	41.66	0.00	4.33	0.51	0.00	1.18	35.14	0.00	5.20	89.46
		PWC	-6.28	1.42	12.75	-12.23	1.48	1.16	-7.84	4.19	6.68	-3.16	0.30	30.46
		$2C_R$	-10.81	1.58	10.45	-21.05	0.38	0.68	-11.44	1.78	5.40	-7.70	2.99	25.32
		3C	-13.55	0.57	6.82	-32.22	0.02	0.07	-18.56	0.23	1.59	-9.28	1.65	18.82
		4C	-14.11	0.42	7.02	-33.25	0.01	0.05	-19.64	0.09	0.81	-9.78	1.29	16.24
	SWS	RS	0.00	3.09	29.89	0.00	4.14	-0.04	0.00	1.27	27.99	0.00	3.40	61.83
		PWC	-3.10	1.20	16.45	-11.13	2.25	2.16	-1.07	2.65	10.92	-2.12	0.55	36.32
		$2C_R$	-9.48	1.71	9.42	-19.68	0.47	1.02	-9.59	1.71	3.41	-6.62	3.48	23.87
		3C	-12.02	0.69	6.99	-32.88	0.02	0.06	-12.70	0.51	2.93	-8.72	1.74	18.01
		4C	-13.34	0.38	6.41	-34.23	0.01	0.04	-13.49	0.21	1.36	-10.27	0.99	17.85
	SWO	RS	0.00	3.35	38.09	0.00	3.15	3.45	0.00	2.17	38.08	0.00	4.45	72.87
		PWC	-2.79	1.98	27.94	-10.63	2.04	10.41	-6.46	1.93	9.98	1.03	5.41	63.52
		$2C_R$	-9.28	1.25	19.87	-19.60	0.47	0.99	-11.55	1.45	10.80	-5.63	1.97	47.89
		3C	-11.68	0.97	15.25	-27.66	0.05	0.27	-13.54	0.88	7.50	-8.01	2.27	38.04
		4C	-12.87	0.62	9.61	-29.67	0.02	0.21	-16.38	0.28	3.84	-8.83	1.76	24.80
Spherical Wave	SWC	RS	0.00	3.88	41.66	0.00	4.33	0.51	0.00	1.18	35.14	0.00	5.20	89.46
		PWC	-4.27	0.24	11.59	-9.43	8.86	-0.01	-4.30	2.73	5.52	-1.95	4.00	29.29
		$2C_R$	-9.02	0.59	11.07	-17.44	0.77	1.25	-10.58	0.63	5.00	-5.67	0.36	27.01
		3C	-12.34	0.61	6.86	-27.34	0.26	-0.40	-16.92	0.13	3.08	-8.15	1.58	17.92
		4C	-16.46	0.25	2.70	-36.34	0.05	-0.04	-21.53	0.09	0.10	-12.17	0.64	8.06
	SWS	RS	0.00	3.09	29.89	0.00	4.14	-0.04	0.00	1.27	27.99	0.00	3.40	61.83
		PWC	-2.49	0.98	14.96	-12.00	8.49	0.67	-0.73	0.73	9.42	-1.06	4.44	34.83
		$2C_R$	-10.75	0.70	11.16	-20.44	0.88	-0.30	-11.28	0.33	3.02	-7.71	0.90	30.79
		3C	-12.84	0.60	5.76	-27.04	0.31	-0.38	-13.85	0.16	1.26	-9.48	1.45	16.42
		4C	-15.54	0.25	2.35	-35.38	0.10	-0.02	-15.14	0.19	0.50	-12.77	0.48	6.57
	SWO	RS	0.00	3.35	38.09	0.00	3.15	3.45	0.00	2.17	38.08	0.00	4.45	72.87
		PWC	-2.49	4.41	27.07	-10.80	6.68	9.54	-6.81	0.35	9.11	1.47	8.44	62.64
		$2C_R$	-9.55	0.58	18.82	-18.98	0.89	-0.31	-12.52	0.55	9.79	-5.78	0.34	47.05
		3C	-12.15	0.92	16.87	-27.25	0.20	0.14	-14.05	0.65	7.76	-8.49	2.16	42.76
		4C	-13.41	0.68	9.40	-29.46	0.16	0.13	-16.88	0.33	3.82	-9.39	1.71	24.27

targets in the range of $10-40$ Hz. The plane wave method showed a similar level of effectiveness for both targets in all ranges. Both the two-channel (rear) and three-channel approaches showed a drop in effectiveness between achieving the plane wave target and spherical wave target in the SWC layout.

For the MSV data in Table A.6, almost all methods of control are below the acceptable threshold of 35 dB. The exceptions are the plane wave cancellation method, the two-channel(rear) and the three-channel in the SWO layout with both targets in the band $70-100$ Hz, which are above 35 dB. Again from Table A.6, it is clear that, again, the

four-channel approach consistently gave MSV values which were always below the 15 dB threshold, denoting a good system.

Comparing the $\bar{\varepsilon}$ values in the full frequency band between Table A.6 and Table A.3, the short wide room offered less effective control for the majority of methods. The improvement in the SWS also meant the plane wave cancellation method in this stereo layout offered effective control over the raw signal, whereas in LTS, the plane wave cancellation method was worse than the raw signal. The rest of the data in Table A.6 followed the same pattern of being worse than the long thin room except for the stereo layout where it is mixed.

All the data in this set of studies again agreed with the data already seen in the chapter, namely that in order of effectiveness, the four-channel, three-channel and two-channel (rear) approaches were more effective than the plane wave cancellation method. Although the change in layout did change the effectiveness of the methods, and the change in the room orientation clearly had a large effect, nothing has arisen from this study to propose a limitation of the least-squares approach.

A.4 Short wide room zones

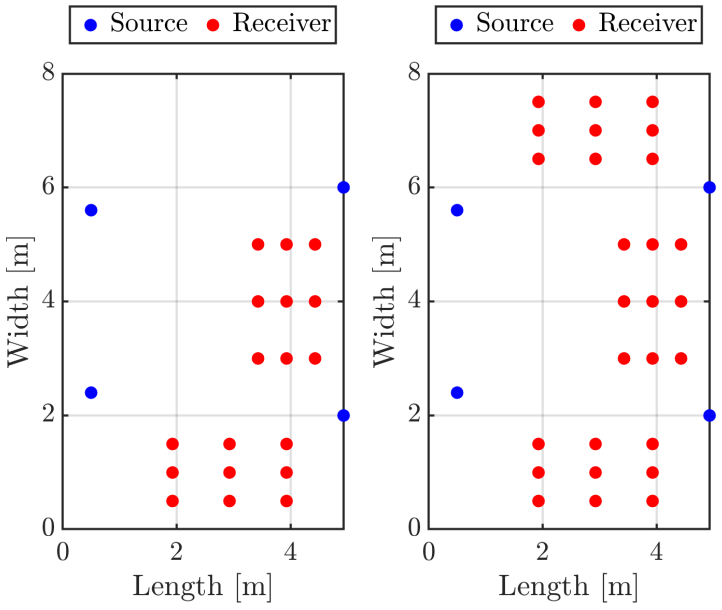
The dimensions of the short wide room allowed more listening areas. Previous studies looked at the effectiveness of the control over a single listening area while changing the front source locations. Here, the stereo layout was used, but the effect of the added second and third listening areas was investigated. The locations of the sources are given in Table A.5 as SWS, and the first listening zone receiver locations are given in Table A.4, listening zone two in Table A.7 and listening zone three in Table A.8. Figure A.7 shows a graphical representation of the two layouts in the test. These were compared with the SWS data from the previous section.

Layout 1: **SWD** (Plot A.7a) – dual listening areas. This also shows the effectiveness of an asymmetrical layout.

Layout 2: **SWT** (Plot A.7b) – triple listening areas.

TABLE A.7: Zone two receiver locations.

Receiver No.	Receiver Location (x m, y m, z m)	Receiver No.	Receiver Location (x m, y m, z m)	Receiver No.	Receiver Location (x m, y m, z m)
r_{10}	(1.93, 0.50, 1.48)	r_{13}	(2.93, 0.50, 1.48)	r_{16}	(3.93, 0.50, 1.48)
r_{11}	(1.93, 1.00, 1.48)	r_{14}	(2.93, 1.00, 1.48)	r_{17}	(3.93, 1.00, 1.48)
r_{12}	(1.93, 1.50, 1.48)	r_{15}	(2.93, 1.50, 1.48)	r_{18}	(3.93, 1.50, 1.48)



(a) SWD (b) SWT.

FIGURE A.7: Short wide room listening zone layouts.

TABLE A.8: Zone three receiver locations.

Receiver No.	Receiver Location (x m, y m, z m)	Receiver No.	Receiver Location (x m, y m, z m)	Receiver No.	Receiver Location (x m, y m, z m)
r_{19}	(1.93, 6.50, 1.48)	r_{22}	(2.93, 6.50, 1.48)	r_{25}	(3.93, 6.50, 1.48)
r_{20}	(1.93, 7.00, 1.48)	r_{23}	(2.93, 7.00, 1.48)	r_{26}	(3.93, 7.00, 1.48)
r_{21}	(1.93, 7.50, 1.48)	r_{24}	(2.93, 7.50, 1.48)	r_{27}	(3.93, 7.50, 1.48)

In this section, Table A.9 shows the band-limited $\bar{\varepsilon}$, $|\text{RMS}_\Delta|$ from the target, and MSV data for the STO, STS and STC layouts with both plane and spherical wave target sound fields for all methods in the test.

In Table A.9, there is a trend in the $\bar{\varepsilon}$ values for the least-squares approaches. The effectiveness dropped as the number of areas of control increased, whereas the effectiveness of the plane wave cancellation method did not drop with more areas of control.

Figure A.8 and Figure A.9 plot the band-limited $\bar{\varepsilon}$ values from Table A.9 for the plane and spherical wave target sound fields, respectively. These plots clearly show the aforementioned drop in the effectiveness of the least-squares approaches as the number of areas of control increased. In all cases, the least-squares approaches were still more effective than the plane wave cancellation method. As asserted, the four-channel approach should always be as good as or better than the plane wave cancellation method. With more receiver points, the least-squares filter averaged for the whole environment, meaning that local issues that may occur in a small number of measurements are less likely to be addressed.

TABLE A.9: Band-limited $\bar{\varepsilon}$ [dB], RMS [dB] and MSV over a fixed listening area for the SWS, SWD and SWT layouts with both plane and spherical wave target sound fields used.

Target	Layout	Method	10Hz $> f \geq$ 100Hz			10Hz $> f \geq$ 40Hz			40Hz $> f \geq$ 70Hz			70Hz $> f \geq$ 100Hz		
			$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV	$\bar{\varepsilon}$	$\lvert RMS_\Delta \rvert$	MSV
Plane Wave	SWS	RS	0.00	3.09	29.89	0.00	4.14	-0.04	0.00	1.27	27.99	0.00	3.40	61.83
		PWC	-3.10	1.20	16.45	-11.13	2.25	2.16	-1.07	2.65	10.92	-2.12	0.55	36.32
		$2C_R$	-9.48	1.71	9.42	-19.68	0.47	1.02	-9.59	1.71	3.41	-6.62	3.48	23.87
		3C	-12.02	0.69	6.99	-32.88	0.02	0.06	-12.70	0.51	2.93	-8.72	1.74	18.01
		4C	-13.34	0.38	6.41	-34.23	0.01	0.04	-13.49	0.21	1.36	-10.27	0.99	17.85
	SWD	RS	0.00	2.58	29.72	0.00	2.77	3.18	0.00	2.14	27.20	0.00	2.51	58.90
		PWC	-4.65	0.81	21.36	-12.37	1.16	4.58	-4.96	2.94	19.10	-1.85	0.84	40.46
		$2C_R$	-8.03	1.77	13.13	-18.80	0.47	1.08	-8.08	2.11	9.66	-5.18	3.15	28.69
		3C	-9.14	1.63	12.39	-26.13	0.04	0.32	-9.21	2.02	10.01	-6.13	3.57	26.87
		4C	-11.30	0.69	8.41	-32.27	0.01	0.09	-14.41	0.27	3.12	-7.31	2.07	22.05
	SWT	RS	0.00	2.49	27.90	0.00	2.30	2.30	0.00	3.19	25.11	0.00	3.65	56.41
		PWC	-4.81	1.69	20.77	-12.90	0.85	4.14	-5.69	3.04	18.12	-1.73	3.93	40.12
		$2C_R$	-7.11	1.87	13.41	-17.93	0.75	0.64	-6.68	1.91	14.73	-4.51	3.34	24.92
		3C	-7.91	1.68	15.08	-24.07	0.09	0.60	-7.68	1.78	18.43	-5.07	4.02	26.27
		4C	-8.48	1.31	14.04	-24.95	0.05	0.53	-8.46	1.14	12.53	-5.53	3.36	29.12
Spherical Wave	SWS	RS	0.00	3.09	29.89	0.00	4.14	-0.04	0.00	1.27	27.99	0.00	3.40	61.83
		PWC	-2.49	0.98	14.96	-12.00	8.49	0.67	-0.73	0.73	9.42	-1.06	4.44	34.83
		$2C_R$	-10.75	0.70	11.16	-20.44	0.88	-0.30	-11.28	0.33	3.02	-7.71	0.90	30.79
		3C	-12.84	0.60	5.76	-27.04	0.31	-0.38	-13.85	0.16	1.26	-9.48	1.45	16.42
		4C	-15.54	0.25	2.35	-35.38	0.10	-0.02	-15.14	0.19	0.50	-12.77	0.48	6.57
	SWD	RS	0.00	2.58	29.72	0.00	2.77	3.18	0.00	2.14	27.20	0.00	2.51	58.90
		PWC	-2.93	2.22	15.61	-9.41	2.09	-1.16	-3.24	1.78	13.35	-0.26	3.87	34.71
		$2C_R$	-8.66	0.62	11.15	-15.67	1.06	-1.20	-8.38	0.40	9.74	-6.25	2.00	24.96
		3C	-11.72	0.94	6.43	-21.86	0.41	-0.02	-11.33	1.24	0.35	-9.14	1.23	18.97
		4C	-12.22	1.19	5.20	-23.92	0.33	-1.13	-11.82	1.23	-0.71	-9.58	2.24	17.46
	SWT	RS	0.00	2.49	27.90	0.00	2.30	2.30	0.00	3.19	25.11	0.00	3.65	56.41
		PWC	-0.71	1.40	12.81	-4.34	3.86	-3.83	-0.33	2.70	10.16	0.99	3.11	32.16
		$2C_R$	-4.79	1.53	18.86	-9.35	2.20	14.41	-3.08	1.95	14.20	-4.09	3.22	28.05
		3C	-8.91	0.96	15.46	-18.90	1.39	0.11	-7.75	1.02	14.10	-6.87	0.53	32.23
		4C	-9.38	1.21	14.14	-21.01	0.71	1.11	-7.97	1.61	12.01	-7.47	1.92	29.36

In Table A.9, the trend in $\bar{\varepsilon}$ is repeated across all the data values, with the plane wave cancellation method being constant (or improving), and for the least-squares approaches, the effectiveness dropping as the number of zones increased. This raised the question about how effective a least-squares approach would be across the whole room, and this was investigated in Section 4.5.2.

The results are summarised in Section 4.4.

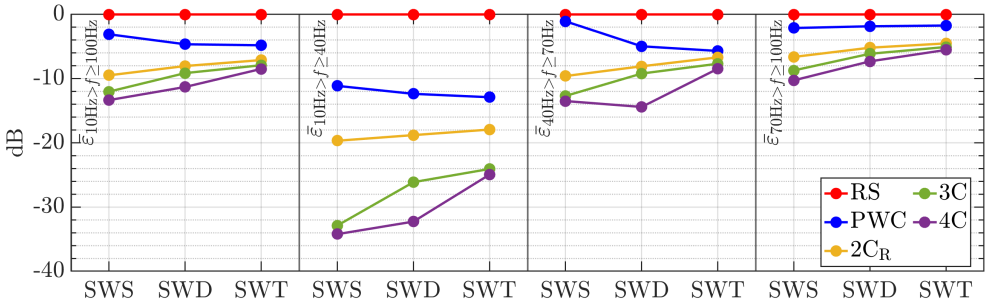


FIGURE A.8: Band-limited $\bar{\varepsilon}$ [dB] over the fixed listening area for the SWS, SWD and SWT layouts with plane wave target sound fields; showing the raw signal (—●—RS), the plane wave cancellation method (—●—PWC), two-channel (rear) approach (- ● -$2C_R$), three-channel approach (—●—3C) and four-channel approach (—●—4C).

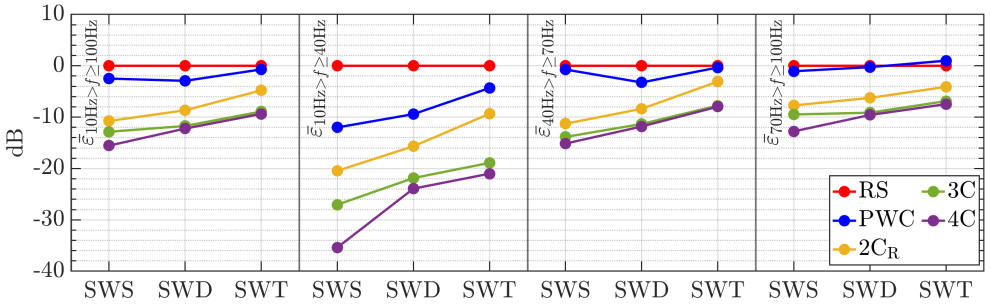


FIGURE A.9: Band-limited $\bar{\varepsilon}$ [dB] over the fixed listening area for the SWS, SWD and SWT layouts with spherical wave target sound fields; showing the raw signal (—●—RS), the plane wave cancellation method (—●—PWC), two-channel (rear) approach (- ● -$2C_R$), three-channel approach (—●—3C) and four-channel approach (—●—4C).

Appendix B

Testing the source location: Spherical wave target sound field

In this appendix, the study from Section 4.5.3 was re-run with the target sound field over the listening areas as a spherical wave, as defined in Equation 4.35. The data in each figure show two histograms for the spherical wave target when using all the defined source locations for both the four-channel approach and the plane wave cancellation method. The minimum, maximum and average values for each set of data are highlighted.

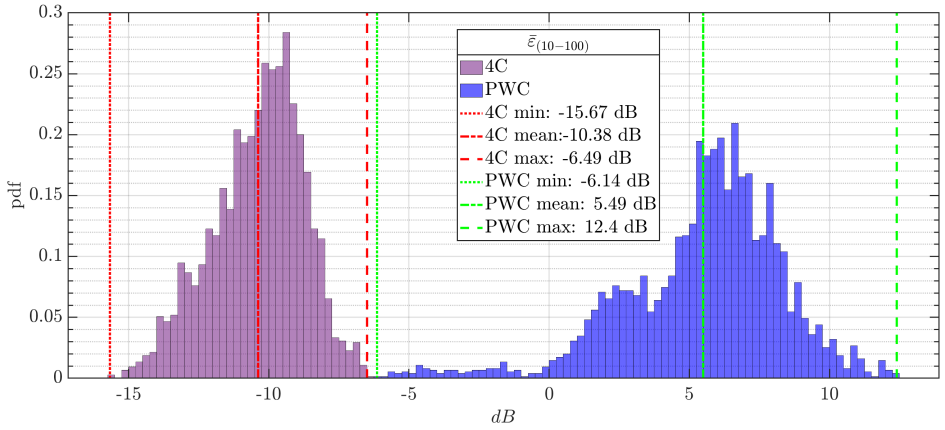


FIGURE B.1: Histograms of $\bar{\varepsilon}$ for all defined source locations with a spherical wave target sound field over the listening area, with the plane wave cancellation method (PWC) and the four-channel approach (4C).

In Figure B.1, the two histograms show the $\bar{\varepsilon}$ in the frequency band. The effectiveness of the four-channel approach had a range of 9.18 dB from –15.67 to –6.49 dB with a mean of –10.38 dB, whereas in Figure 4.33, the four-channel approach had a range of 10.27 dB from –16.37 to –6.1 dB with a mean of –9.7 dB. The performance of the four-channel approach was relatively consistent with both targets. At its best, it was slightly more

effective at achieving the plane wave target, but at its worst, it was slightly more effective at achieving the spherical wave target. On average, the four-channel approach was 0.68 dB more effective at achieving the spherical wave target. Obversely, the plane wave cancellation method was on average 0.61 dB less effective at achieving the spherical wave target. Importantly, the four-channel approach was still no worse than the raw signal at the 0 dB reference for any of the 3,000 possible source combinations, as it always offered an average improvement of at least 6.49 dB.

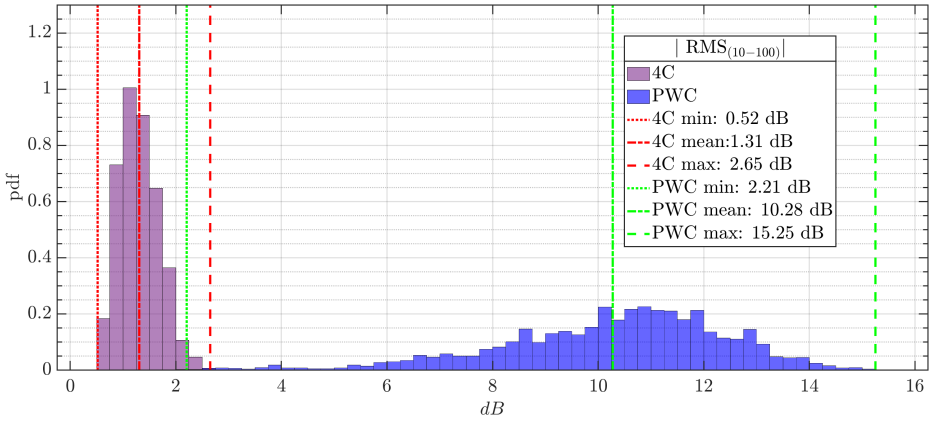


FIGURE B.2: Histograms of $|\text{RMS}_\Delta|$ for all defined source locations with a spherical wave target sound field over the listening area, with the plane wave cancellation method (PWC) and the four-channel approach (4C).

In Figure B.2, the two histograms show the $|\text{RMS}_\Delta|$ from the target in the frequency band. The four-channel approach outperformed the plane wave cancellation method, with the lowest plane wave cancellation value at 2.21 dB, which was 0.44 dB better than the worst four-channel approach value and 0.9 dB worse than the mean four-channel approach value. For these data, the four-channel approach had a range of 2.13 dB from 0.52 to 2.65 dB with a mean of 1.31 dB, whereas in Figure 4.34, the four-channel approach had a range of 2.22 dB from 0.4 to 2.62 dB with a mean of –1.18 dB. The difference here was minimal, but it showed that for this measure, the four-channel approach was very slightly more effective at achieving the plane wave target.

Figure B.3 shows the two histograms of the MSV from the target in the frequency band. The majority of the data for the plane wave cancellation method had an MSV value greater than the 35 dB threshold, which meant that the variation between the receivers was too high for the system to have any acceptable value, as stated in Section 2.3. As for the four-channel method, the inverse is true; the majority of the data is below the 35 dB threshold, and with a mean of 17.28 dB, most of the locations are, therefore, either in the low quadrant of the acceptable range ($\leq$ 20dB) or below the 15 dB threshold and therefore denoted as good.

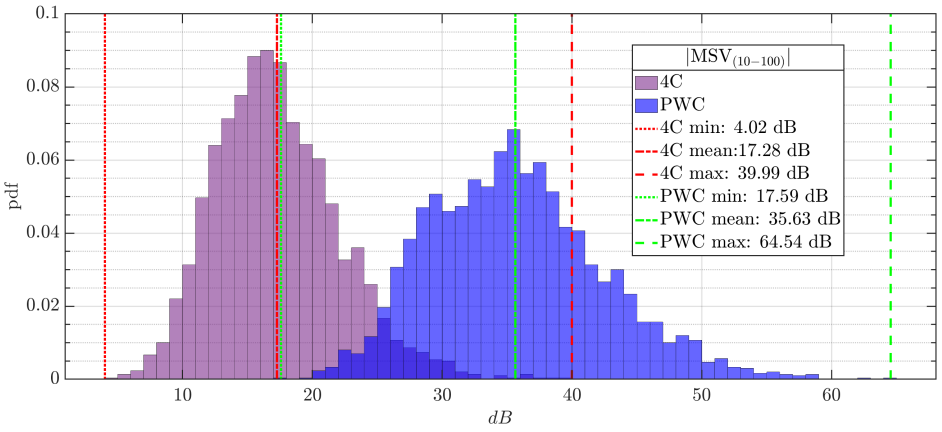


FIGURE B.3: Histograms of MSV| for all defined source locations with a spherical wave target sound field over the listening area, with the plane wave cancellation method (PWC) and the four-channel approach (4C).

References

861v8 DS v2.7. 861v8 reference digital surround controller. white paper, meridian audio., Cambridge, UK, May 2012. URL http://www.meridian-support.com/datasheets/861v8-A4.pdf.

AES. Multichannel surround sound systems and operations. Technical Document AESTD1001.1.01-10, AES, AES, 132 East 43rd Street, Suite 405, NY, NY 10017, 2001.

G. Alkin. *Sound recording & reproduction*. Focal, 3rd edition, 1996.

J. B. Allen and D. A. Berkley. Image method for efficiently simulating small-room acoustics. *The Journal of the Acoustical Society of America*, 65(4):943–950, April 1979. . URL http://dx.doi.org/10.1121/1.382599.

G. M. Amdahl. Validity of the single processor approach to achieving large scale computing capabilities. In *Proceedings of the April 18-20, 1967, Spring Joint Computer Conference*, AFIPS '67 (Spring), page 483–485, New York, NY, USA, 1967. Association for Computing Machinery. ISBN 9781450378956. . URL https://doi.org/10.1145/1465482.1465560.

F. Antonacci, J. Filos, M. R. P Thomas, E. A. P. Habets, S Augusto, P. Naylor, and S. Tubaro. Inference of room geometry from acoustic impulse responses. *IEEE Transactions on Audio, Speech, and Language Processing*, 20:2683–2695, December 2012. . URL https://doi.org/10.1109/TASL.2012.2210877.

D Aprea, F. Antonacci, A. Sarti, and S. Tubaro. Acoustic reconstruction of the geometry of an environment through acquisition of a controlled emission. In *2009 17th European Signal Processing Conference*, pages 710–714, Glasgow, UK, August 2009. IEEE. ISBN 978-161-7388-76-7. URL https://ieeexplore.ieee.org/document/7077681.

T. O. Bell. An investigation into the automated digital correction of the surround sound images produced from non-optimum loudspeaker arrangements. Master's thesis, University of Hertfordshire, 2011.

T. O. Bell and M. Baratelli. A Comparison Between Plane Wave Cancellation and a Least-Squares Approach for the Zonal Control of Modes in a Rectangular Room. In *Audio Engineering Society Convention 148*, Online, June 2020. AES, 132 East 43rd

Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=20759.

T. O. Bell and F. M. Fazi. An Investigation into the Location and Number of Microphone Measurements Necessary for Efficient Active Control of Low-Frequency Sound Fields in Listening Rooms. In *Audio Engineering Society Convention 147*, New York, USA, October 2019. AES, 132 East 43rd Street, Suite 405, NY, NY 10017.

L. Beranek and T. Mellow. *Acoustics*. Academic Press, 2nd edition, 2019.

A. J. Berkhout. A Holographic Approach to Acoustic Control. *Journal of the Audio Engineering Society*, 36(12):977–995, December 1988. URL http://www.aes.org/e-lib/browse.cfm?elib=5117.

A. J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *The Journal of the Acoustical Society of America*, 93(5):2764–2778, 1993. . URL http://dx.doi.org/10.1121/1.405852.

S. Bharitkar. System And Method For Automatic Room Acoustic Correction. U.S. Patent 2003/0235318 A1, 2003.

S. Bharitkar and C. Kyriakakis. A cluster centroid method for room response equalization at multiple locations. In *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, pages 55–58. IEEE, February 2001. ISBN 0-7803-7126-7. . URL https://doi.org/10.1109/ASPAA.2001.969541.

S. Bharitkar and C. Kyriakakis. System and method for automatic room acoustic correction in multi-channel audio environments. *The Journal of the Acoustical Society of America*, 129:1668, March 2011. .

S. Bharitkar, P. Hilmes, and C. Kyriakakis. Robustness of spatial averaging equalization methods: a statistical approach. In *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, 2002.*, volume 1, pages 184–187. IEEE, November 2002. URL https://ieeexplore.ieee.org/document/1197172.

S. Bharitkar, P. Hilmes, and C. Kyriakakis. Sensitivity of multichannel room equalization to listener position. In *International Conference on Multimedia and Expo. ICME '03.*, pages I – 721. IEEE, October 2003. . URL https://doi.org/10.1109/ICME.2003.1221019.

J Blauert. *Spatial Hearing: The Psychophysics of Human Sound Localization*. The MIT Press, 10 1996. ISBN 9780262268684. . URL https://doi.org/10.7551/mitpress/6391.001.0001.

M. Bodley. High directivity microphone array. *The Journal of the Acoustical Society of America*, 113:2957–2957, January 2003. . URL https://doi.org/10.1121/1.2227659.

P. Bruel. *Sound insulation and room acoustics*. Chapman and Hall, 2-6 Boundary Row, London, SE1 8HN, U.K., 1st edition, 1951.

P. Budd, M. Smith, and K. Robertson. Method for optimizing the performance of a loudspeaker to compensate for low frequency room modes. G.B. Patent GB2519676A;, April 2015.

M. A. Burgess and W. A. Utley. Reverberation times in British living rooms. *Applied Acoustics*, 18(5):369–380, 1985. ISSN 0003-682X. . URL https://www.sciencedirect.com/science/article/pii/0003682X85900556.

S. Cecchi, A. Carini, and S. Spors. Room Response Equalization—A Review. *Applied Sciences*, 8(1):16, December 2017. ISSN 2076-3417. . URL http://dx.doi.org/10.3390/app8010016.

A. Celestinos and S. B. Nielsen. Multi-Source Low Frequency Room Simulation Using Finite Difference Time Domain Approximations. In *Audio Engineering Society Convention 117*, San Francisco, USA, October 2004. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=12921.

A. Celestinos and S. B. Nielsen. Optimizing Placement and Equalization of Multiple Low Frequency Loudspeakers in Rooms. In *Audio Engineering Society Convention 119*, New York, USA, October 2005. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=13307.

A. Celestinos and S. B. Nielsen. Low Frequency Sound Field Enhancement System for Rectangular Rooms using Multiple Low Frequency Loudspeakers. In *Audio Engineering Society Convention 120*, Paris, France, May 2006. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=13492.

A. Celestinos and S. B. Nielsen. Low-Frequency Loudspeaker–Room Simulation Using Finite Differences in the Time Domain—Part 1: Analysis. *Journal of the Audio Engineering Society*, 56(10):772–786, 2008a. URL http://www.aes.org/e-lib/browse.cfm?elib=14636.

A. Celestinos and S. B. Nielsen. Controlled Acoustic Bass System (CABS) A Method to Achieve Uniform Sound Field Distribution at Low Frequencies in Rectangular Rooms. *Journal of the Audio Engineering Society*, 56(11):915–931, November 2008b. URL http://www.aes.org/e-lib/browse.cfm?elib=14642.

J. Cheer. *Active Control of the Acoustic Environment in an Automobile Cabin*. PhD thesis, University of Southampton, 1 2012. URL https://eprints.soton.ac.uk/348944/.

D. R. Conant and W. A. Beyer. Generalized pythagorean theorem. *The American Mathematical Monthly*, 81(3):262–265, 1974. URL http://www.jstor.org/stable/2319528.

T. J. Cox, P. D'antonio, and M. R. Avis. Room sizing and optimization at low frequencies. *Journal of the audio engineering society*, 52(6):640–651, june 2004. .

M. Crocco, A. Trucco, and A. D. Bue. Uncalibrated 3d room reconstruction from sound. *ArXiv*, abs/1606.06258, 2016.

N. Devos. Reconstruction of room geometry through echolocation for surround audio applications. Master's thesis, Ghent University, 2019.

I. Dokmanić, R. Parhizkar, A. Walther, Y. M. Lu, and M. Vetterli. Acoustic echoes reveal room shape. *Proceedings of the National Academy of Sciences*, 110(30):12186–12191, 2013. ISSN 0027-8424. URL https://www.pnas.org/content/110/30/12186.

S. J. Elliott. *Signal Processing for Active Control*. Elsevier Science, 2000.

S. J. Elliott and P. A. Nelson. Multiple-point equalization in a room using adaptive digital filters. *AES: Journal of the Audio Engineering Society*, 37(11):899–907, 1989. ISSN 00047554.

F. A. Everest. *The Master Handbook of Acoustics*. TAB Books, 1989. ISBN 9780830690961. URL https://books.google.co.uk/books?id=nQRRAAAAMAAJ.

F. Fahy and J. Walker. *Fundamentals of Noise and Vibration*. E & FN Spon, 11 New Fetter Lane, London, EC4P 4EE, 1st edition, 1998.

A. Farina. Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. In *Audio Engineering Society Convention 108*, Paris, France, 2000. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=10211.

B. Fazenda, M. Wankling, J. Hargreaves, L. Elmer, and J. Hirst. Subjective preference of modal control methods in listening rooms. *Journal of the Audio Engineering Society*, 60 (5):338–349, May 2012. URL http://www.aes.org/e-lib/browse.cfm?elib=16324.

J. Filos, E. Habets, and P Naylor. A two-step approach to blindly infer room geometries. In *Intl. Workshop Acoust. Echo Noise Control (IWAENC)*, Tel Aviv, Israel, August 2010. IWAENC.

J. Filos, A. Canclini, M. Thomas, F. Antonacci, A. Sarti, and P. Naylor. Robust Inference of Room Geometry from Acoustic Measurements Using the Hough Transform. In *19th European Signal Processing Conference*, Barcelona, Spain, April 2011. IEEE.

E. Georganti, T. May, S. van de Par, A. Harma, and J. Mourjopoulos. Speaker Distance Detection Using a Single Microphone. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(7):1949–1961, 2011. . URL https://doi.org/10.1109/TASL.2011.2104953.

Y. Haneda, S. Makino, and Y. Kaneda. Common Acoustical Pole and Zero Modeling of Room Transfer Functions. *Speech and Audio Processing, IEEE Transactions on*, 2:320–328, April 1994. . URL https://doi.org/10.1109/89.279281.

Y Haneda, S Makino, and Y Kaneda. Multiple-point equalization of room transfer functions by using common acoustical poles. *IEEE Transactions on Speech and Audio Processing*, 5(4):325–333, 1997. URL https://doi.org/10.1109/89.593306.

P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems*. SIAM, 1998. ISBN 9780898714036.

P. C. Hansen. *Computational Inverse Problems in Electrocardiology*, chapter The L-Curve and Its Use in the Numerical Treatment of Inverse Problems, pages 119–142. WIT-press, 2001.

S. Hardy, M. E. Johnson, S. K. Porter, and J. H. Sheerin. Loudspeaker with reduced audio coloration caused by reflections from a surface. U.S. Patent 10,015,584B, July 2018.

W. Hoeg, L. Christensen, and .R Walker. Subjective assessment of audio quality - the means and methods within the ebu. Ebu technical review, European Broadcasting Union, Geneva, Switzerland, 1997.

IEC:60268-13. Sound system equipment - part 13: Listening tests on loudspeakers. Standard, International Electrotechnical Commission, Geneva, CH, November 1998. URL https://webstore.iec.ch/publication/1213.

ITU-R:BS.775-3. Multichannel stereophonic sound system with and without accompanying picture. Standard, International Telecommunication Union, Geneva, CH, August 2012. URL https://www.itu.int/rec/R-REC-BS.775/en.

G. M. Jackson and H. G. Leventhall. The acoustics of domestic rooms. *Applied Acoustics*, 5(4):265–277, 1972. ISSN 0003-682X. . URL https://www.sciencedirect.com/science/article/pii/0003682X72900308.

K. Jambrošić, M. Horvat, and H. Domitrović. Reverberation time measuring methods. *The Journal of the Acoustical Society of America*, 123(5):3617–3617, May 2008. . URL https://doi.org/10.1121/1.2934829.

M. Johansson. On room correction and equalization of sound systems. In *Audio Engineering Society Convention 123*, New York, USA, October 2007. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL https://www.dirac.com/wp-content/uploads/2021/09/On-equalization-filters.pdf.

M. Karjalainen, A. Makivirta, P. Antsalo, and V. Valimaki. Low-frequency modal equalization of loudspeaker-room responses. In *Audio Engineering Society Convention 111*,

New York, USA, November 2001. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=9849.

L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sander. *Fundamentals of acoustics*. John Wiley & Sons, 3rd edition, 1982.

O. Kirkeby and P. A. Nelson. Reproduction of plane wave sound fields. *The Journal of the Acoustical Society of America*, 94(5):2992–3000, 1993. . URL https://doi.org/10.1121/1.407330.

O. Kirkeby and P. A. Nelson. Digital Filter Design for Inversion Problems in Sound Reproduction. *Journal of the Audio Engineering Society*, 47(7/8):583–595, 1999. URL http://www.aes.org/e-lib/browse.cfm?elib=12098.

O. Kirkeby, P. A. Nelson, H. Hamada, and F. Orduna-Bustamante. Local sound field reproduction using digital signal processing. *The Journal of the Acoustical Society of America*, 100(3), September 1996. URL https://resource.isvr.soton.ac.uk/staff/pubs/PubPDFs/Pub3634.pdf.

H. Kuttruff. *Room Acoustics*. CRC Press, Taylor \& Francis Group, 6000 Broken Sound Parkway NW Suite 300, Boca Raton, FL 33487-2742, 6st edition, 2017.

M. M. Louden. Dimension-ratios of rectangular rooms with good distribution of eigentones. *Acta Acustica united with Acustica*, 24(2), 1971. ISSN 1610-1928. URL https://www.ingentaconnect.com/content/dav/aaua/1971/00000024/00000002/art00007.

Y. Luan and F. Jacobsen. A method of measuring the Green's function in an enclosure. *The Journal of the Acoustical Society of America*, 123(6):4044–4046, 2008. . URL https://doi.org/10.1121/1.2917804.

W. Ma and X. Liu. Phased Microphone Array for Sound Source Localization with Deep Learning. *Aerospace Systems*, 2, 2 2018. . URL https://doi.org/10.1007/s42401-019-00026-w.

W. Ma, B. Huan, C. Zhang, and X. Liu. Beamforming of phased microphone array for rotating sound source localization. *Journal of Sound and Vibration*, 467:115064, 11 2019. . URL https://doi.org/10.1016/j.jsv.2019.115064.

R. Magalotti and V. Cardinali. A simulation test bench for decay times in room acoustics. In *2018 COMSOL Conference*, Lausanne, Italy, 2018. COMSOL,100 District Avenue, Burlington, MA 01803, USA. URL https://www.comsol.com/paper/a-simulation-test-bench-for-decay-times-in-room-acoustics-66091.

R. Magalotti and V. Cardinali. Building FEM low frequency room models through modal decay time measurements. In *23rd International Congrees on acoustics*, Aachen,

Germany, 2019. URL http://publications.rwth-aachen.de/record/769718/files/769718.pdf.

R. Magalotti and D. Ponteggia. Use of wavelet transform for the computation of modal decay times in rooms. In *Audio Engineering Society Convention 147*, New York, USA, October 2019. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=20608.

V. Magalotti, R. Cardinali. Modal Decay Times in Ducts and Rooms. In *Audio Engineering Society Convention 146*, Dublin Ireland, March 2019. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=20320.

J. Martínez-Mesa, D. A. González-Chica, J. L. Bastos, R. R. Bonamigo, and R. P. Duquia. Sample size: how many participants do i need in my research? *Anais brasileiros de dermatologia*, 89(4):609–615, 2014. URL https://doi.org/10.1590/abd1806-4841.20143705.

M. Meissner. A novel method for determining optimum dimension ratios for small rectangular rooms. *Archives of Acoustics*, vol. 43(No 2), 2018. URL http://journals.pan.pl/Content/107269/AoA_122369.pdf.

M. Miyoshi and Y. Kaneda. Inverse control of room acoustics using multiple loudspeakers and/Or microphones. In *ICASSP '86. IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 11, pages 917–920, Tokyo, Japan, 1986. IEEE. URL https://ieeexplore.ieee.org/document/1168847.

M. Miyoshi and Y. Kaneda. Inverse filtering of room acoustics. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(2):145–152, February 1988. URL https://ieeexplore.ieee.org/document/1509.

A. H. Moore, M. Brookes, and P. A. Naylor. Room geometry estimation from a single channel acoustic impulse response. In *21st European Signal Processing Conference (EUSIPCO 2013)*, pages 1–5, Marrakech, Morocco, September 2013. IEEE. URL https://ieeexplore.ieee.org/document/6811439.

G. E. Moore. Cramming more components onto integrated circuits. *Electronics magazine*, 38(8):114, April 1965. URL https://ieeexplore.ieee.org/document/4785860.

P. M. Morse and K. U. Ingard. *Theoretical acoustics*. McGraw-Hill, New York, 1968.

J. N. Mourjopoulos. Digital Equalization of Room Acoustics. *Journal of the Audio Engineering Society*, 42(11):884–900, 1994. URL http://www.aes.org/e-lib/browse.cfm?elib=6921.

S. T. Neely and J. B. Allen. Invertibility of a room impulse response. *The Journal of the Acoustical Society of America*, 66(1):165–169, 1979. . URL https://doi.org/10.1121/1.383069.

P. A. Nelson and S. J. Elliott. Least squares approximations to exact multiple point sound reproduction. *Institute of Sound and Vibration Research memorandum 683*, 1988.

P. A. Nelson and S. J. Elliott. *Acitve Control of Sound*. Academic Press inc, 24/28 Oval Road, London NW1 7DX, 1st edition, 1992.

S. Norcross, M. Bouchard, and G. Soulodre. Inverse filtering design using a minimal-phase target function from regularization. In *Audio Engineering Society Convention 121*, pages 5–8, San Francisco, USA, 2006. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=13763.

A. Noxon. Room Acousitcs and Low Frequency damping. In *Audio Engineering Society Convention 81*, Los Angeles, USA, 11 1986. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL https://www.acousticsciences.com/wp-content/uploads/1986/11/aes-1986-room-acoustics.pdf.

F. Olivieri, M. Shin, F. M. Fazi, P. A. Nelson, and P. Otto. Loudspeaker array processing for multi-zone audio reproduction based on analytical and measured electroacoustical transfer functions. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control - Engineering and Perception*, Sep 2013. URL http://www.aes.org/e-lib/browse.cfm?elib=16897.

A. O'Neill. Average living room sizes in newly-built houses in britain each decade since the 1930s (in square meters. *Statista*, October 2020. URL https://www.statista.com/statistics/1056003/average-living-room-sizes-new-british-houses-1930-2020/.

A. V. Oppenheim, A. S. Willsky, and I. T. Young. *Signals and Systems*. Prentice-Hall, New York, 1982.

J. Pedersen. A Method and System For Adapting A Loudspeaker To A Listening Position In A Room. U.S. Patent US 8,094,826 B2, January 2012.

A. D. Pierce. *Acoustics: an introduction to its physical principles and applications*. New York: McGraw-Hill, 1st edition, 1981.

L. Remaggi, P Jackson, P Coleman, and W Wang. Room boundary estimation from acoustic room impulse responses. In *2014 Sensor Signal Processing for Defence (SSPD)*, Edinburgh, UK, September 2014. IEEE. URL https://ieeexplore.ieee.org/document/6943328.

F. Ribeiro, D. Ba, C. Zhang, and D. Florencio. Turning enemies into friends: Using reflections to improve Sound Source Localization. In *2010 IEEE International Conference on Multimedia and Expo, ICME 2010*, Singapore, September 2010. IEEE. .

Y. Rui, D. Florencio, W. Lam, and J. Su. Sound source localization for circular arrays of directional microphones. In *IEEE International Conference on Acoustics, Speech, and*

Signal Processing., volume 3, pages iii/93 – iii/96 Vol. 3, Philadelphia, PA, USA, May 2005. IEEE. ISBN 0-7803-8874-7. . URL https://doi.org/10.1109/ICASSP.2005.1415654.

F. Rumsey and T. McCormick. *Sound and recording: an introduction*. Focal, 4nd edition, 2002.

W. C. Sabine. *Collected Papers on Acoustics*. Cambridge : Harvard University Press, 1st edition, 1922.

A. O. Santillán. Spatially extended sound equalization in rectangular rooms. *The Journal of the Acoustical Society of America*, 110(4):1989–1997, 2001. . URL http://dx.doi.org/10.1121/1.1401740.

A. O. Santillán, C. S. Pedersen, and M. Lydolf. Experimental implementation of a low-frequency global sound equalization method based on free field propagation. *The Journal of the Acoustical Society of America*, 68(10):1063–1085, 2007. URL https://doi.org/10.1016/j.apacoust.2006.05.010.

M. R. Schroeder. The "Schroeder frequency" revisited. *The Journal of the Acoustical Society of America*, 99:3240–3241, 1996. URL https://doi.org/10.1121/1.414868.

J. Scott and B. Dragovic. Audio Location: Accurate Low-Cost Location Sensing. In H.-W. Gellersen, R. Want, and A. Schmidt, editors, *Pervasive Computing*, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

P. Seetharaman and S. Tarzia. The Hand Clap as an Impulse Source for Measuring Room Acoustics, April 2012. URL http://www.aes.org/e-lib/browse.cfm?elib=16223.

M. Shin, F. M. Fazi, P. A. Nelson, and F. Hirono. Controlled sound field with a dual layer loudspeaker array. *Journal of Sound and Vibration*, 333(16):3794–3817, 2014. ISSN 0022-460X. . URL http://www.sciencedirect.com/science/article/pii/S0022460X14002260.

M. F. Simón Gálvez, S. J. Elliott, and J. Cheer. A superdirective array of phase shift sources. *The Journal of the Acoustical Society of America*, 132(2):746–756, 2012. . URL https://doi.org/10.1121/1.4733556.

S. Smith. *Digital signal processing: A practical guide for engineers and scientists*. Newnes, Burlington MA, 2003.

I. Sokolnikoff. *Mathematics of physics and modern engineering*. McGraw-Hill, Inc, New York, 1966.

S. Spors and J. Ahrens. Spatial Sampling Artifacts of Wave Field Synthesis for the Reproduction of Virtual Point Sources. In *Audio Engineering Society Convention 126*,

Munich, germany, May 2009. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=14940.

K. A. Stroud and J. B. Booth. *Engineering Mathematics*. Palgrave Macmillan, January 2001. ISBN 978-0333919392.

S. P. Tarzia, P. A. Dinda, R. P. Dick, and G. Memik. Indoor Localization without Infrastructure Using the Acoustic Background Spectrum. In *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services*, MobiSys '11, pages 155–168, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450306430. . URL https://doi.org/10.1145/1999995.2000011.

S. Tervo and T. Tossavainen. 3D room geometry estimation from measured impulse responses. In *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on*, pages 513–516, Kyoto, Japan, March 2012. IEEE. ISBN 978-1-4673-0045-2. . URL https://doi.org/10.1109/ICASSP.2012.6287929.

S. Tervo, J. Pätynen, and T. Lokki. Acoustic reflection path tracing using a highly directional loudspeaker. In *2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 245–248, New Paltz, NY, USA, 2009. IEEE. .

H. Tokuno, O. Kirkeby, P. A. Nelson, and H. Hamada. Inverse filter of sound reproduction systems using regularization. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E80-A(5):809–820, 1997. ISSN 09168508. URL http://resource.isvr.soton.ac.uk/staff/pubs/PubPDFs/Pub3731.pdf.

F. E. Toole. Loudspeaker measurements and their relationship to listener preferences: part 1. *Journal of the Audio Engineering Society*, 34(4):227–235, april 1986.

J. Vanderkooy. Multi-source room equalization: reducing room resonances. Paper presented at 123rd AES Convention. In *Audio Engineering Society Convention 123*, New York, USA, October 2007. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=14320.

J. Vanderkooy. The low-frequency acoustic centre: Measurement, theory and application. In *Audio Engineering Society Convention 128*, London, UK, May 2010. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=15289.

J. Vanderkooy. New Thoughts on Active Acoustic Absorbers. In *Audio Engineering Society Convention 131*, New York, USA, October 2011. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=15984.

J. Vanderkooy and M. Rousseau. Active Acoustic Absorbers Revisited. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control - Engineering and Perception*, Guildford, UK, September 2013. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=16909.

D. Ward and T. Abhayapala. Reproduction of a Plane-Wave Sound Field Using an Array of Loudspeakers. *IEEE Transaction on Speech and Audio Processing*, 9:697–707, September 2001. . URL https://ieeexplore.ieee.org/document/943347.

T. Welti. Investigation of bonello criteria for use in small room acoustics. In *Audio Engineering Society Convention 127*, New York, USA, October 2009. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=15044.

T Welti. Optimal configurations for subwoofers in rooms considering seat to seat variation and low frequency efficiency. In *Audio Engineering Society Convention 133*, San Francisco, USA, October 2012. AES, 132 East 43rd Street, Suite 405, NY, NY 10017. URL http://www.aes.org/e-lib/browse.cfm?elib=16490.

T. Welti and A. Devantier. Low-frequency optimization using multiple subwoofers. *Journal of the audio engineering society*, 54:347–364, May 2006. URL http://www.aes.org/e-lib/browse.cfm?elib=13680.

J. Weng and K. Guentchev. Three-dimensional sound localization from a compact non-coplanar array of microphones using tree-based learning. *The Journal of the Acoustical Society of America*, 110:310–323, August 2001. . URL https://doi.org/10.1121/1.1377290.

E. Williams. *Fourier acoustics: sound radiation and nearfield acoustical holography*. Academic Press, London, 1999.

X. Wu, S. Jin, L. I. Yibo, and P. Zhang. Sound source localization based on microphone arrays. *Chinese Journal of Sensors and Actuators*, 23:682–686, May 2010.

C. Zhang, D. Florencio, D. Ba, and Z. Zhang. Maximum Likelihood Sound Source Localization and Beamforming for Directional Microphone Arrays in Distributed Meetings. *IEEE Transactions on Multimedia*, 10:538–548, April 2008. . URL https://doi.org/10.1109/TMM.2008.917406.