

Analysis, Modeling and Wide-Area Spatiotemporal Control of Low-Frequency Sound Reproduction

Adam J. Hill



A thesis submitted for the degree of Doctor of Philosophy

School of Computer Science & Electronic Engineering

University of Essex

January 2012

Summary

This research aims to develop a low-frequency response control methodology capable of delivering a consistent spectral and temporal response over a wide listening area. Low-frequency room acoustics are naturally plagued by *room-modes*, a result of standing waves at frequencies with wavelengths that are integer multiples of one or more room dimension. The standing wave pattern is different for each modal frequency, causing a complicated sound field exhibiting a highly position-dependent frequency response.

Enhanced systems are investigated with multiple *degrees of freedom* (independently-controllable sound radiating sources) to provide adequate low-frequency response control. The proposed solution, termed a *chameleon subwoofer array* or *CSA*, adopts the most advantageous aspects of existing room-mode correction methodologies while emphasizing efficiency and practicality.

Multiple degrees of freedom are ideally achieved by employing what is designated a *hybrid subwoofer*, which provides four orthogonal degrees of freedom configured within a modest-sized enclosure. The CSA software algorithm integrates both objective and subjective measures to address listener preferences including the possibility of individual real-time control.

CSAs and existing techniques are evaluated within a novel acoustical modeling system (*FDTD simulation toolbox*) developed to meet the requirements of this research. Extensive virtual development of CSAs has led to experimentation using a prototype hybrid subwoofer. The resulting performance is in line with the simulations, whereby variance across a wide listening area is reduced by over 50% with only four degrees of freedom.

A supplemental novel correction algorithm addresses correction issues at select narrow frequency bands. These frequencies are filtered from the signal and replaced using virtual bass to maintain all aural information, a psychoacoustical effect giving the impression of low-frequency. Virtual bass is synthesized using an original hybrid approach combining two mainstream synthesis procedures while suppressing each method's inherent weaknesses. This algorithm is demonstrated to improve CSA output efficiency while maintaining acceptable subjective performance.

Table of contents

SUMMARY.....	1
ACKNOWLEDGEMENTS.....	6
1 INTRODUCTION	7
1.1 Problem statement.....	7
1.2 Proposed solution.....	8
1.3 Thesis organization	8
1.4 Claims for originality.....	10
2 LOW-FREQUENCY ROOM ACOUSTICS	12
2.1 Overview.....	12
2.2 Room-modes	15
2.2.1 Definition of the low-frequency spectral range	16
2.2.2 Theoretical modal contributions.....	17
2.2.3 Spatial variance	18
2.2.4 Magnitude deviation	19
2.2.5 Mean output level	19
2.3 Influential factors	20
2.3.1 Surface absorption and reactance	20
2.3.2 The Waterhouse effect.....	20
2.3.3 Helmholtz resonance.....	21
2.3.4 Temperature and humidity	21
2.4 Measurement techniques	22
2.4.1 Maximum length sequences (MLS)	22
2.4.2 Tone bursts.....	25
2.5 Chapter summary	26
3 LOW-FREQUENCY MODELING	27
3.1 Simulation methods.....	27
3.1.1 Image source	27
3.1.2 Ray tracing	28
3.1.3 Finite element method (FEM) and boundary element method (BEM).....	29
3.1.4 Hybrid methods.....	30
3.2 Finite-difference time-domain theory	30
3.2.1 Structure	30
3.2.2 Boundary conditions	32
3.2.3 Advanced boundary conditions.....	33
3.2.4 Stability	33
3.2.5 Non-rectangular FDTD simulation.....	34
3.2.6 Non-rectangular and room obstacle grid masking.....	36
3.3 Simulation toolbox overview	39
3.3.1 Core simulation code.....	39
3.3.2 MATLAB graphical user interface (GUI) programming.....	40
3.4 Simulation validation.....	42
3.4.1 Theoretical modes.....	42
3.4.2 Lab measurements	44

3.4.3	Previously published results.....	45
3.4.4	Discussion	47
3.5	Chapter summary	48
4 CONVENTIONAL ROOM-MODE CORRECTION.....		50
4.1	Passive correction.....	50
4.1.1	Room dimensions	50
4.1.2	Absorption	53
4.1.3	Single subwoofer placement.....	54
4.1.4	Multiple subwoofers	57
4.1.5	Passive correction conclusions.....	58
4.2	Active correction	59
4.2.1	Static single-point equalization	59
4.2.2	Adaptive single-point equalization.....	61
4.2.3	Multiple-point equalization.....	62
4.2.4	Polar pattern control	63
4.2.5	Absorption	68
4.2.6	Radiation resistance	68
4.2.7	Ambisonics-style equalization.....	70
4.2.8	Active correction conclusions.....	71
4.3	Chapter summary	72
5 CHAMELEON SUBWOOFER ARRAYS.....		73
5.1	Theoretical concept	74
5.1.1	Degrees of freedom.....	74
5.1.2	Hybrid sources.....	75
5.1.3	Mathematical structure.....	75
5.2	System practicality issues	77
5.2.1	Dipole efficiency	77
5.2.2	Lower correction limit	78
5.2.3	Upper correction limit	79
5.2.4	Windowing	80
5.3	Correction performance.....	84
5.3.1	One-unit CSA (four degrees of freedom)	84
5.3.2	Four-unit CSA (sixteen degrees of freedom).....	90
5.3.3	Vertical correction range.....	93
5.4	Performance versus array configuration.....	95
5.4.1	Array layout performance	96
5.4.2	Source spacing.....	99
5.4.3	Source position offset	101
5.4.4	Single versus multiple drive-unit subwoofers	102
5.4.5	Layout recommendations	105
5.5	Additional target points and pseudo-inverse filtering.....	106
5.5.1	Variable listening area, fixed target point density	106
5.5.2	Fixed listening area, variable target point density.....	107
5.5.3	Pseudo-inverse filtering conclusions.....	108
5.6	CSA prototype development	108
5.6.1	Omnidirectional subwoofer CSA experimentation	108
5.6.2	Prototype hybrid subwoofer development.....	114
5.6.3	Hybrid subwoofer CSA experimentation.....	115
5.6.4	Discussion	117

5.7	Performance comparisons	117
5.8	Chapter summary	119
6	VIRTUAL BASS SYSTEMS.....	120
6.1	Principles of virtual bass synthesis	120
6.2	Time-domain implementation – nonlinear devices.....	121
6.2.1	Core possessing procedure and code.....	122
6.2.2	Discussion	125
6.3	Frequency-domain implementation – phase vocoders.....	126
6.3.1	Core processing procedure and code.....	127
6.3.2	Discussion	133
6.4	Hybrid implementation.....	134
6.4.1	Transient content detector	134
6.4.2	Subjective evaluation	140
6.4.3	Discussion	142
6.5	Room-mode correction	142
6.5.1	Subjective evaluation procedure.....	143
6.5.2	Subjective evaluation results.....	144
6.5.3	Chameleon subwoofer array applications	145
6.5.4	Discussion	148
6.6	Chapter summary	148
7	EXTENDED APPLICATIONS OF CHAMELEON SUBWOOFER ARRAYS	150
7.1	Location-specific equalization	150
7.1.1	Low-frequency control procedure	150
7.1.2	Response manipulation examples.....	152
7.1.3	Discussion	154
7.2	Live sound	155
7.2.1	Chameleon subwoofer array implementation	155
7.2.2	Pseudo-inverse filtering for live sound CSAs.....	158
7.2.3	Discussion	160
7.3	Conventional home theater	160
7.3.1	Example scenario.....	161
7.3.2	Discussion	163
7.4	Chapter summary	163
8	CONCLUSIONS.....	165
8.1	Critical assessment	165
8.1.1	FDTD simulation toolbox	165
8.1.2	Chameleon subwoofer arrays	166
8.1.3	Virtual bass systems	167
8.2	Future work.....	168
8.2.1	FDTD simulation toolbox	168
8.2.2	Chameleon subwoofer arrays	169
8.2.3	Virtual bass systems	170
8.3	Closing remarks	170
REFERENCES.....		171

APPENDIX A – FDTD SIMULATION EQUATIONS AND MASK SET	177
A.1 Two-dimensional FDTD simulation equations.....	177
A.1.1 Non-boundary particle velocity.....	177
A.1.2 All pressure values.....	177
A.1.3 Characteristic wall impedance & particle velocity constants	178
A.1.4 Boundary particle velocity (x-dimension)	178
A.1.5 Boundary particle velocity (y-dimension)	178
A.2 Three-dimensional FDTD simulation equations	179
A.2.1 Non-boundary particle velocity.....	179
A.2.2 All pressure values.....	179
A.2.3 Characteristic wall impedance & particle velocity constants	179
A.2.4 Boundary particle velocity (x-dimension)	180
A.2.5 Boundary particle velocity (y-dimension)	180
A.2.6 Boundary particle velocity (z-dimension).....	180
A.3 Non-rectangular topology example mask set	181
A.3.1 User-defined pressure element grid mask (pressure updates)	181
A.3.2 Non-boundary particle velocity element grid masks (velocity updates)	181
A.3.3 Boundary particle velocity element grid masks (velocity updates)	182
A.3.4 Pressure element grid masks (non-boundary velocity updates)	183
A.3.5 Pressure element grid masks (boundary velocity updates)	184
A.3.6 Velocity element grid masks (pressure updates)	185
APPENDIX B – FDTD SIMULATION TOOLBOX EXAMPLES.....	186
B.1 Toolbox features – configuration	186
B.1.1 Basic topology definition.....	186
B.1.2 Spatial configuration	188
B.1.3 Source setup	189
B.1.4 Signal generation.....	189
B.1.5 Virtual listener setup	192
B.2 Toolbox features – simulation.....	192
B.2.1 Animations.....	192
B.2.2 Final spatio-pressure distribution plots	194
B.2.3 Auralization.....	195
B.3 Toolbox features – analysis.....	196
B.3.1 Room layout	196
B.3.2 Theoretical mode calculator.....	197
B.3.3 Time domain & frequency domain analysis	199
B.3.4 Spectrogram	199
B.3.5 Tone bursts	200
B.4 Worked examples.....	202
B.4.1 Modal buildup animations	202
B.4.2 Non-rectangular topologies	203
B.4.3 Network of rooms	204
B.4.4 Subwoofer polar pattern control visualizations.....	206
B.4.5 Loudspeaker enclosure modeling	207
APPENDIX C – JOURNAL PUBLICATIONS	209
APPENDIX D – CONFERENCE PAPERS.....	210

Acknowledgements

First and foremost, I must thank my supervisor, Professor Malcolm Hawksford, for his never-ending support and motivation over the past three years. Malcolm not only provided his immense wealth of knowledge towards guiding the subject matter of this project, but also imparted valuable lessons concerning academic writing and presentation methods which helped me significantly improve my communication skills.

Additionally, I would like to thank Gary Geaves and Jon Moore at Bowers & Wilkins for their initial inspiration for this project and their continued support through extremely useful feedback and the donation of hardware for the system prototype.

I must express my gratitude to my family for their continued encouragement during my education, even as my studies have taken me far away overseas. The value of growing up in a home that encouraged and supported pursuing your dreams cannot be understated.

At Essex, Adrian Boland-Thoms provided his expertise for the fabrication of the prototype subwoofer and Adam Balkan assisted with key measurements during various experiments. Also, my former labmates, Daisuke Koya and Ephraim Gower, provided encouragement as well as useful feedback through numerous discussions and all my friends outside the lab supplied, as their title suggests, good reason to escape the lab from time to time.

Lastly, I would be remiss if I did not acknowledge Mars, Inc. for their delicious product, peanut M&M's, which have largely fuelled me during this research. I cannot give a precise figure concerning the quantity consumed, but it is certainly not insignificant, so I thank you (seriously).

1 Introduction

The primary objective of sound reproduction has historically been to disseminate recorded aural information. Although the primary focus has changed little over the years, system objectives have expanded considerably. No longer is the average listener content by simply hearing program material, but now expects to be enveloped by a virtual soundscape, either accurately replicating the recorded environment or synthesizing a sound field otherwise impossible in the real world. These heightened expectations translate into greater sound reproduction system requirements, necessitating hardware and software capable of accurately manipulating a given acoustical environment to match that of the program material and to give users a consistently positive listening experience.

An ongoing task within the acoustics and audio engineering community focuses on the effective interfacing between the electro-mechanical sound reproduction system (loudspeakers, power amplifiers, processing and playback units) and the acoustical environment. Closed acoustical spaces strongly color the transmitted temporal and spectral response of a reproduction system, largely due to wall reflections, and consequently result in position-dependent listening experiences where some (or many) listeners may not experience the intended soundscape.

In the high-frequency range, problems largely lie with stereo/surround imaging which can be smeared if an individual is significantly offset from the central listening area or if there are strong wall reflections. While this problem is largely solvable by repositioning the main loudspeakers (for stereo systems) or adding supplementary loudspeakers (for surround systems), solving problems inherent to the low-frequency band is not as straightforward.

1.1 Problem statement

Low-frequency room acoustics are shaped by what are referred to as room-modes. When a frequency's wavelength is an integer multiple of one or more room dimension a standing wave pattern results, hence a room-mode at that frequency. This occurrence causes certain areas within the space to receive strong amplitudes at the modal frequency while other areas receive nearly nothing. The standing wave pattern is different for all modal frequencies, causing a complicated variable low-frequency response across the listening area.

There exists a wealth of solutions targeting variance minimization. Some call for alterations to the physical configuration of the room and/or sound reproduction system. Others implement signal processing routines based on room measurements. In a few cases, low-frequency polar pattern control is proposed to minimally excite the most problematic room-modes, generally by utilizing multiple drive-units within each subwoofer and adhering to a suggested source placement to minimize source-to-room coupling.

The aim of this research is to develop a method for wide-area low-frequency spatiotemporal response control by combining aspects of previously proven techniques as well as establishing novel system components, both in the hardware and software domains, to overcome deficiencies characteristic of the various approaches. Broadly, the intended outcome calls for a low-frequency sound reproduction system capable of delivering a consistent spectral and temporal response across a large listening area while simultaneously considering the subjective requirements of individuals within the space. Both the hardware and software should be practically realizable with emphasis placed on keeping both the required quantity of sound reproduction units and the computational requirements to a minimum while still achieving acceptable results.

1.2 Proposed solution

The centerpiece of the proposed low-frequency room-mode correction technique is the concept of exploiting multiple degrees of freedom within a system construct. Degrees of freedom are defined in this work as relating to independently controllable radiating sources within the sound reproduction space. For example, a conventional low-frequency source (subwoofer) contains just one drive-unit which contributes a single degree of freedom towards system correction. This inevitably limits the possibilities for sound manipulation, especially in an attempt to limit frequency response variance among listeners.

The presented research explores enhanced systems having multiple degrees of freedom to provide adequate low-frequency response control. This is achieved either by creating arrays of conventional, single drive-unit subwoofers or by employing what is designated a *hybrid subwoofer* that has been developed specifically for this project and which provides up to four orthogonal degrees of freedom configured within a modest-sized enclosure.

The proposed implementation uses a software-based algorithm to help achieve the desired results. The newly-established method results from an extensive study of existing digital signal processing (DSP) methods utilized for similar purposes. It adopts the most advantageous aspects of each routine while placing constraints/bounds to ensure efficient and consistent system operation. The algorithm integrates both objective and subjective (psychoacoustical) measures that can address listener preferences including the possibility of customized real-time individual control. The hybrid loudspeaker and software package are complementary and form a system termed a *chameleon subwoofer array* or *CSA*. The CSA DSP is shown to be compatible with existing hardware structures to achieve similar results as when paired with hybrid subwoofers.

1.3 Thesis organization

This thesis is structured to reflect the interdependence of various subject areas considered germane to this project where Fig. 1.1 depicts a high-level overview explaining the interconnectivity of the core topics and chapters. The research is presented in chronological order to lead readers through a logical progression of thought commencing with a generalized literature review which becomes progressively more specific with relevant original research and results presented, as needed, until the objectives are adequately met approaching the culmination of the work.

A discussion on low-frequency room acoustics (Chapter 2) bestows the required background necessary to comprehend the targeted acoustical issues of this project (expressed in general terms in Section 1.1). This commences with a discussion on room acoustics including various metrics for acoustical response characterization followed by an introduction to room-modes, the crux of the problem in small-room low-frequency acoustics. The chapter concludes with a presentation of useful metrics to objectively classify a low-frequency response among numerous listening locations and also commonplace techniques for measuring said values, including maximum length sequences (MLS) and tone bursts.

An investigation into low-frequency modeling is undertaken in Chapter 3 to provide a toolset for virtual prototyping. Techniques are considered from the geometrical, statistical and hybrid categories examining each method's advantages/disadvantages based upon findings in published literature. Finite-difference time-domain (FDTD) modeling is chosen for this project as it is shown to meet the simulation requirements and is straightforward to implement.

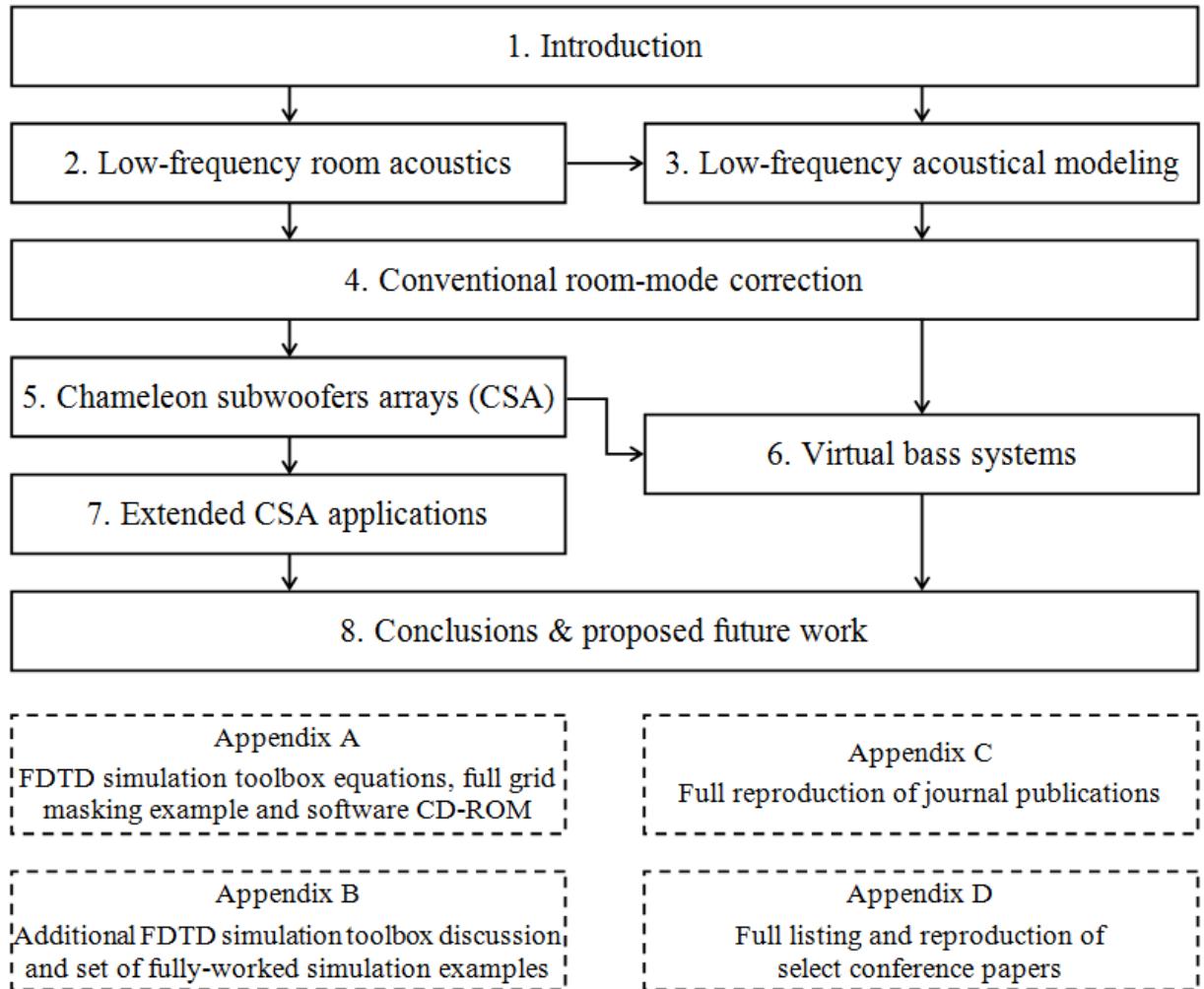


Fig. 1.1 Organization of the main body of the thesis

Following a survey of the available simulation software, it was determined that this research required the development of a bespoke acoustical modeling toolbox to allow for adequate simulation flexibility and data analysis. This led to the development of the FDTD simulation toolbox which is described in the latter sections of Chapter 3 with further detail (including a full set of modeling equations and additional worked examples) presented in Appendices A and B. The discussion focusing on the simulation toolbox includes result-validation using comparisons of simulated responses to theoretical, measured and other published simulations to verify the toolbox's accuracy.

With a software package capable of modeling a wide variety of physical topologies containing various low-frequency sound reproduction systems, the research critically examines conventional room-mode correction techniques (Chapter 4). The considered methods fall within two designated categories, namely passive and active correction. In this work, passive correction is defined as control by just the physical configuration of loudspeakers within the listening space, whereas active correction includes signal processing. As a benchmark most common correction techniques are examined using the FDTD toolbox to demonstrate their capabilities and corresponding drawbacks.

The analysis in Chapter 4 provides the room-mode correction framework, leading directly to the development of a novel routine (both software and hardware based), termed a *chameleon*

subwoofer array (CSA), detailed extensively in Chapter 5. This chapter begins by emphasizing the focus placed on *degrees of freedom* and how the idea naturally leads to the hybrid subwoofer (with multiple degrees of freedom) approach. As this system is aligned using a direct calculation procedure, a detailed discussion on system practicality is included that examines the most appropriate frequency range of operation as well as limitations related to source and measurement-point layout. The chapter concludes with a set of directives for implementation of such a system, including practical results obtained from the system prototype development.

As discussed in Chapter 5, it is shown that narrow frequency bands can exist that impose unrealistic demands on the reproduction system. A novel solution to this problem is proposed in Chapter 6, whereby the psychoacoustical phenomenon of virtual bass can be targeted to replace the physical reproduction of a limited number of more problematic narrow frequency bands. This procedure can subjectively maintain transmission of all aural data within the program material while easing physical reproduction requirements. This approach is applicable to any correction system, or can even be used as a standalone system, although particular emphasis is placed here on its applicability to CSAs.

Chapters 4 – 6 concentrate predominantly on achieving an even spatiotemporal response over a wide-area within small-sized listening rooms. The inherent flexibility of a CSA lends itself to broader applications, three of which are discussed in Chapter 7. First, the CSA algorithm allows each listener the ability to manipulate their localized frequency response without affecting surrounding users, where this can be performed in real-time due to the CSA direct calculation procedure. Second, this localized correction application is explored in the context of large-scale live-sound reinforcement where systems are required to deliver an even distribution of low-frequency energy over a wide audience area while reducing stage levels to protect the performers' hearing. It is shown that the CSA algorithm can be built into the system processors without the need for new hardware. Lastly, the DSP-only CSA approach is applied to small-room home theater scenarios by adapting the procedure to match a 5.1 surround sound system.

The thesis concludes in Chapter 8 with a summary of the research presented, with emphasis placed on the results obtained together with a reinforcement of claims for originality. The chapter contains suggestions for future work principally directed towards focusing system development towards a marketable product, either in home theater or live sound reinforcement. A set of appendices follows Chapter 8 containing the previously mentioned additional discussion together with worked examples from the FDTD simulation toolbox (Appendices A & B) and reproductions of journal publications (Appendix C) and select conference papers (Appendix D).

1.4 Claims for originality

The original aspects of this research fall into three core subject areas: low-frequency acoustical modeling, chameleon subwoofer arrays and virtual bass (although all are targeted towards chameleon subwoofer array implementation).

The original aspects of **low-frequency acoustical modeling** are:

- 1) The *FDTD simulation toolbox* presents a unique collection of acoustical modeling and data analysis routines configured within an easy to use package. Most included components are previously known (aside from points 2 – 4 in this list), although containing all elements within a single piece of software is believed unique.

- 2) A mask-based implementation of non-rectangular topology modeling is developed allowing for simulation of rooms of any shape. The procedure lends itself to modeling room obstacles, a network of rooms or even loudspeaker enclosures.
- 3) Several routines are included within the toolbox including: single/multiple source placement optimization, room dimension ratio optimization, live sound subwoofer placement optimization and a live sound subwoofer cluster configuration tool.
- 4) The toolbox has a useful set of visualization capabilities allowing for inspection of sound propagation during simulation (either in two or three dimensions) and detailed spatio-pressure plots for post-simulation analysis. Also included is a collection of data plotting tools, specifically tailored for certain test signals such as tone bursts.

The original aspects of **chameleon subwoofer arrays** are:

- 1) The design of the hybrid subwoofer, which ideally can consist of six drive-units (one per cubic-enclosure face) creating four orthogonal source components which correspondingly contributes four degrees of freedom towards system correction.
- 2) A direct-calculation routine for low-frequency correction allowing for precise real-time adjustable sound-field manipulation with constraints in place to avoid driver over-excursion or correction outside the effective low-frequency band.
- 3) The combination of signal processing and multiple loudspeakers (conventional and/or hybrid) to obtain accurate low-frequency sound reproduction over a wide area, exhibiting superior time and frequency domain performance (time domain is regularly ignored in most existing approaches).
- 4) Individualized low-frequency equalization capabilities are included in the correction routine allowing users to adjust their localized frequency response without significantly affecting surrounding users.
- 5) A live-sound low-frequency coverage pattern control is developed, largely based on (4), addressing the need for even sound dispersion across the audience while limiting energy on stage.

The original aspects of **virtual bass** are:

- 1) A hybrid nonlinear device (NLD) and phase vocoder (PV) virtual bass synthesis technique whereby a transient content detector (TCD) monitors the input signal and dynamically weights the two virtual bass synthesis procedures (running in parallel) based on which best serves the signal. The technique is shown to be signal content-insensitive, unlike mainstream NLD and PV virtual bass systems.
- 2) A virtual bass room-mode correction procedure which, in its standalone form, utilizes multiple parametric equalization band-stop filters to remove problematic room-mode frequency bands from physical reproduction and replaces them with virtual bass components tailored to subjectively reinforce the aural data not physically present. The procedure is intended as a supplement to low-frequency correction techniques.

2 Low-frequency room acoustics

Room acoustics is a multifaceted topic within audio engineering drawing extensively on key principles of physics, specifically fluid dynamics. An in depth exploration of room acoustics can highlight how acoustics influences sound reproduction within a variety of spaces ranging from small closed spaces to large outdoor environments.

The fundamental aspect of small-room low-frequency acoustics, the primary focus of this research, involves the physical layout of the acoustical space. This includes room dimensions, surface material properties, subwoofer placement and listener location, among a variety of other factors. Each of these aspects must be considered during the design of a listening room with a sound reproduction system.

This section accentuates previously published work concerning these topics. A mathematical explanation of key low-frequency acoustical attributes is presented followed by an exploration of the critical factors that shape the low-frequency acoustical response within a closed space along with well-established objective measurement techniques used to evaluate a room's acoustical characteristics.

2.1 Overview

All sound within a room originates at one or more source. Sources emit a sound pressure wavefront, which is evident in spatio-pressure plots where the wavefront appears as a line connecting points of equal sound level (Fig. 2.1) [10].

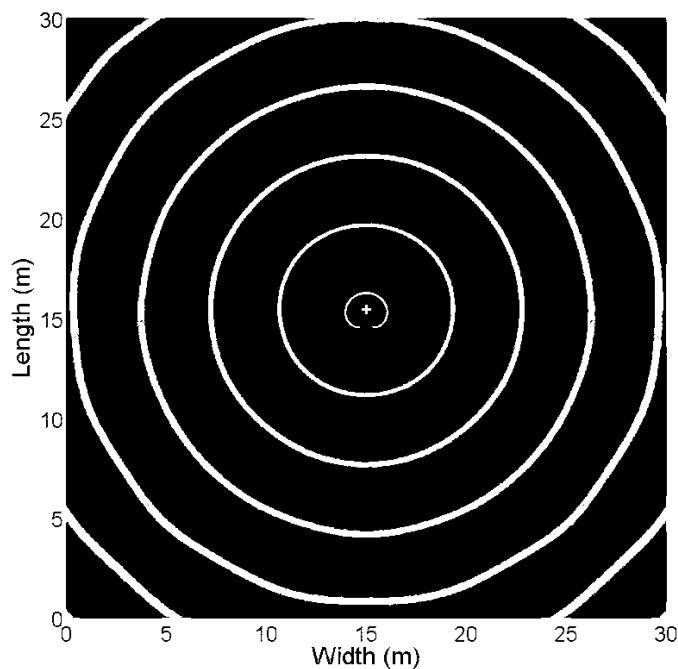


Fig. 2.1 Simulated example of sound pressure wavefronts emitted from a sealed loudspeaker at the center of the diagram

Two metrics commonly used to measure sound at a point in space are sound intensity and sound pressure. Sound intensity is a measure of sound power per unit area in watt per square meter (W/m^2). This is commonly converted to a logarithmic scale where the value is termed sound intensity level, which is calculated from Eq. 2.1 [1].

$$L_I = 10 \log(I/I_0) \quad (2.1)$$

where:	L_I	=	sound intensity level (dB)
	I	=	sound intensity (W/m^2)
	I_0	=	$10^{-12} \text{ W}/\text{m}^2$, reference sound intensity

Sound pressure is a more common metric in practice, as it is easily measured with a pressure microphone. Sound pressure, measured in Pascal (Pa), is proportional to the square root of sound intensity and can similarly be converted to a logarithmic scale with Eq. 2.2; a value known as sound pressure level (SPL) [1]. Both sound intensity level and sound pressure level are measured in decibels (dB).

$$L_{p,free} = 20 \log(p/p_0) \quad (2.2)$$

where:	$L_{p,free}$	=	free-field sound pressure level (dB)
	p	=	sound pressure (Pa)
	p_0	=	$20 \times 10^{-6} \text{ Pa}$, reference sound pressure

Sound intensity is directly proportional to the square of sound pressure ($I \propto p^2$), given that $I = p v$, where v is the particle velocity (m/s) [1]. Note that the coefficient in Eq. 2.2 is 20 rather than 10 due to the sound intensity to sound pressure relationship. Restoring the exponents for the pressure terms in Eq. 2.2 results in a coefficient of 10 (Eq. 2.3).

$$L_{p,free} = 10 \log(p/p_0)^2 \quad (2.3)$$

In a free-field, the wavefront intensity decreases by a factor of four for every doubling of distance from the source. This is known as the inverse square law (Eq. 2.4) [1].

$$I = \frac{P}{4\pi r^2} \quad (2.4)$$

where:	P	=	source power (W)
	r	=	radiation distance (m)

The law operates on the principle that sound radiates equally in all directions, creating a spherical leading edge with a radius corresponding to distance from the source. The amount of instantaneous energy radiated from the source is fixed, covering a surface area of $4\pi r^2$ when the radius of the spherical wavefront is r and expanding to $4\pi(2r)^2 = 16\pi r^2$ when the radius is doubled to $2r$. The quadrupled surface area results in the sound intensity to surface area relationship described by the inverse square law. This corresponds to a halving of sound pressure (-6 dB sound pressure level) for every doubling of radiation distance.

The inverse square law does not necessarily hold true in smaller closed spaces. When wavefronts encounter a surface within a space they are partially absorbed (and transmitted) through the surface and partially reflected back in the space at an angle normal to the angle of incidence (Fig. 2.2) [10].

In the simplest of cases, surfaces are treated as perfectly rigid with frequency-independent absorption. This means that a sound wave encountering a simple surface is absorbed/reflected equally over the entire audio spectrum. In more complex (and realistic) cases, surfaces exhibit frequency-dependent absorption. A sound wave encountering a non-rigid surface experiences different levels of absorption/reflection across the frequency spectrum. In addition, the surface

may resonate at various frequencies, as defined by the material properties, causing additional colorations to the reflections.

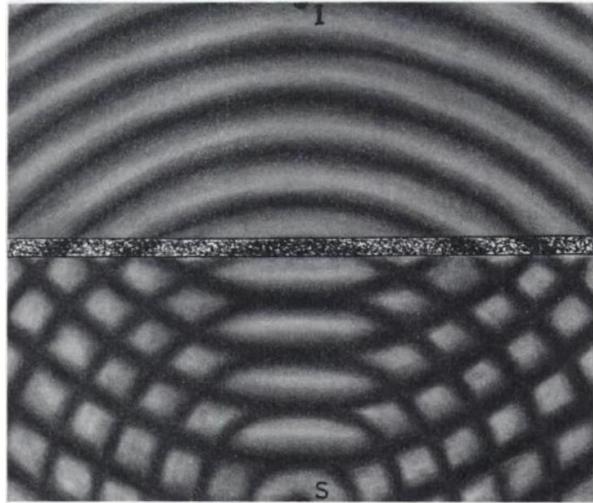


Fig. 2.2 Absorption, reflection and transmission of a sound wave by a wall [10]

These non-anechoic properties distort the direct waveform emitted from the source(s) through interference between the direct and reflected (reverberant) wavefronts. The inverse square law is therefore not accurate in smaller reverberant spaces due to the strong influence of the room on the sound field [1, 2]. In small rooms the overall pressure level is not a critical issue, given sufficient listening distances from the source. In fact, small rooms often exhibit a stronger reverberant field than direct field which is seen with $4/A > Q/4\pi r^2$ in Eq. 2.5, where $4/A$ and $Q/4\pi r^2$ represent the reverberant and direct sound field, respectively [1].

$$L_p = L_w + 10 \log \left(\frac{Q}{4\pi r^2} + \frac{4}{A} \right) \quad (2.5)$$

where:	L_p	=	sound pressure level (dB)
	L_w	=	sound power level from source (dB)
	Q	=	directivity factor
	A	=	total absorption (m^2)

The reverberant sound field contribution, $4/A$, is derived from the fact that in the steady-state the rate of energy absorption by the walls must balance the total power output of the source. If this is not the case (i.e. $A = 0$), the steady state sound pressure level will diverge to infinity. The precise derivation of this component is beyond the scope of this investigation, but can be found in Chapter 6 of [97].

Sound power level is found by comparing the power of a source to a reference power of 10^{-12} W [1] and then converting to a logarithmic scale in decibels (dB). For example, if a sound source produces 5 W of acoustical power, the sound power level is calculated as $L_w = 10\log(5/W_0) \approx 127 \text{ dB}$, where $W_0 = 10^{-12} \text{ W}$.

The directivity factor, Q , of a source is found by taking the ratio of the sound intensity at a distance r in front of the source to the average sound intensity in all directions. An omnidirectional source results in a ratio of 1:1, giving a directivity factor of one. If a source radiates in a hemispherical pattern the ratio is 2:1, giving a directivity factor of 2. A more detailed discussion on source directivity is presented in Chapter 4.

Total absorption, A , is calculated if a room's surface absorptive properties are known where each material in the room has an absorption coefficient, α , ranging from zero (no absorption) to one (full absorption) and a surface area, S , in square meters (Eq. 2.6) [1, 2].

$$A = S\alpha \quad (2.6)$$

As discussed earlier, surface absorption is usually frequency-dependant, so care must be taken when choosing coefficients for absorption calculations in Eq. 2.6. Absorptive characteristics of common materials are published over a number of frequency bands in the audible spectrum and can be used to make appropriate calculations.

2.2 Room-modes

A 60 dB drop in sound pressure level is commonly utilized to characterize a room's reverberation attributes, either by measurement or calculation (Eq. 2.7), as developed by Sabine in the late 19th century [4]. This 60 dB drop generally is reached after a sound wave travels a distance of twenty to thirty times the largest dimension of a room [3].

$$RT_{60} = \frac{c_s V}{A} \quad (2.7)$$

where:	RT_{60}	=	reverberation time (s)
	c_s	=	constant factor, 0.161
	V	=	room volume (m^3)

The reverberation time of a room can vary widely across the frequency spectrum due to the frequency-dependent absorption characteristics of a room. A number of researchers have argued that reverberation time is not a useful metric at low-frequencies due to what are known as room-modes [5-9].

The large number of reflections from a sound source gives rise to standing waves within a room which occur when the half-wavelength of a particular frequency is a simple integer multiple of one or more combination of lengths between parallel surfaces [2]. The frequencies at which these standing waves occur are referred to as eigenfrequencies, or room-modes. Room-modes significantly contribute to the coloration effect a room has on a playback system's response. Theoretical room-mode frequencies are calculated for rectangular rooms using Eq. 2.8 [2]. Of course, room-modes exist in all closed spaces, although non-rectangular spatial distributions are difficult to predict using closed-form solutions.

$$f_m = \frac{c}{2} \sqrt{\left(\frac{\eta_x}{L_x}\right)^2 + \left(\frac{\eta_y}{L_y}\right)^2 + \left(\frac{\eta_z}{L_z}\right)^2} \quad (2.8)$$

where:	f_m	=	theoretical room-modes (Hz)
	c	=	speed of sound in air (m/s)
	η_x, η_y, η_z	=	modal indices (0, 1, 2, ...)
	L_x, L_y, L_z	=	room dimensions (m)

There are three primary varieties of room-modes. The first occurs between a single set of parallel surfaces, termed an axial mode. This is represented in Eq. 2.8 when only one modal index (η_x , η_y or η_z) is non-zero. The second variety of room-mode is called a tangential mode,

occurring between two sets of parallel surfaces with two non-zero modal indices in Eq. 2.8. The last primary type of room-mode is an oblique mode. Oblique modes occur between three sets of parallel surfaces with all modal indices non-zero. Higher order room-modes occur in spaces with more than three sets of parallel surfaces, but can be difficult to calculate and are best predicted using simulation software [2].

The issue encountered (and central to the current research) with room-modes is that the received sound pressure level is largely dependent on a listener's location within the standing wave pattern. This can result in substantially different aural perceptions between adjacent listeners. This phenomenon is discussed later in this chapter.

2.2.1 Definition of the low-frequency spectral range

Although room-modes are present throughout the entire frequency spectrum, they become much less noticeable as frequency increases. The reduced room-mode sensitivity with rising frequency is due to increases in modal spatial and spectral proximity. The human ear discriminates between frequency components using the Basilar membrane in the cochlea [11, 12]. At a certain point in the frequency spectrum, room-modes are sufficiently dense so that the human ear cannot distinguish individual modes due to its limited frequency resolution. In addition, as frequency increases the modal spatial distribution varies over such a short distance that the differences in level are averaged within the auditory processing system and adjacent listeners have similar perceptions (disregarding any subjective biasing or stereo/surround imaging effects) [2].

The transition point most often used to describe the boundary between modal (low-frequency) and diffuse (high-frequency) bands is known as the Schroeder frequency, first presented in 1954 [13] and revised in 1996 [14] by M.R. Schroeder (Eq. 2.9). An alternative form of this equation is presented by Geddes in [15] (Eq. 2.10).

$$f_c = 2000 \sqrt{\frac{RT_{60}}{V}} \quad (2.9)$$

$$f_c = \frac{5000}{\sqrt{V\bar{\alpha}}} \quad (2.10)$$

where: f_c = Schroeder frequency (Hz)
 $\bar{\alpha}$ = mean absorption coefficient in a room

A detailed derivation of the Schroeder frequency (Eq. 2.9) is contained in [9]. The derivation makes an approximation of modal spacing by calculating the derivative of the equation for the theoretical number of room-modes below a given frequency (Eqs. 2.11 & 2.12). The modal spacing approximation (Eq. 2.11) is applied to an equation for the minimum number of modes located within the mean half-width of a resonance within the room (Eq. 2.13) that results in a diffuse field, where distinct room-modes are not noticeable (Eqs. 2.14 & 2.15).

$$N(f) \approx \frac{4\pi V}{3c^3} f_M^3 \quad (2.11)$$

$$\partial f_n = \frac{c^3}{4\pi V f_c^2} \quad (2.12)$$

$$\Delta f_n \approx \frac{cA}{8\pi V} \quad (2.13)$$

$$\Delta f_n > m\delta f_n \quad (2.14)$$

$$f_c^2 > m \frac{2c^2}{A} \quad (2.15)$$

where:	f_M	=	upper limit for modal analysis (Hz)
	$N(f)$	=	number of modes below f_M (Hz)
	δf_n	=	mean modal spacing (Hz)
	Δf_n	=	mean resonance half-width (Hz)
	m	=	scaling constant

An optimal value of $m = 3$ was found experimentally in [16]. With $m = 3$, Eq. 2.15 can be simplified to match Eq. 2.9, the Schroeder frequency equation. The above derivation describes how a sound field is considered diffuse when enough room-modes' spectral resonant curves intersect within an average half-bandwidth of the mean modal room resonance (Eq. 2.14).

2.2.2 Theoretical modal contributions

In the discrete modal range (below the Schroeder frequency) variations in response are easily perceived as a listener moves about a room. This variation is due to fluctuations in room to source coupling. These factors can be calculated for defined source or receiver locations at each room-mode (Eq. 2.16) [7].

$$\psi(x, y, z) = \cos\left(\frac{\eta_x \pi x}{L_x}\right) \cos\left(\frac{\eta_y \pi y}{L_y}\right) \cos\left(\frac{\eta_z \pi z}{L_z}\right) \quad (2.16)$$

where:	ψ	=	modal distribution function
	x, y, z	=	source/receiver location coordinates (m)

The net contribution of a room-mode to the overall room response is calculated using the resulting modal distribution functions from Eq. 2.16 (Eq. 2.17) [7].

$$|p| = \frac{\epsilon_{N,x} \epsilon_{N,y} \epsilon_{N,z} \psi(s) \psi(R)}{\frac{2\omega_N \kappa_N}{\omega} + j \left(\frac{\omega_N^2}{\omega} - \omega \right)} \quad (2.17)$$

where:	$ p $	=	modal sound pressure contribution (Pa)
	$\epsilon_{N,x}, \epsilon_{N,y}, \epsilon_{N,z}$	=	modal order scaling factor for mode N
	$\psi(s)$	=	source distribution function for source s
	$\psi(R)$	=	receiver distribution function for receiver R
	ω_N	=	natural angular frequency of mode N (rad/s)
	ω	=	angular frequency under inspection (rad/s)
	κ_N	=	damping factor of mode N

The modal damping factors are calculated using the total absorption, A , for each room dimension along with room volume and the relevant modal scaling factors. The modal scaling factors used in published research are one for the zero-order mode (0 Hz) and two for all other modes (Eq. 2.18) [7].

$$\kappa_N = \frac{c}{8V} \frac{\epsilon_{N,x}A_x + \epsilon_{N,y}A_y + \epsilon_{N,z}A_z}{2} \quad (2.18)$$

Finally, the overall sound pressure is found for a given frequency with specified source and receiver locations using a summation of each individual modal contribution (Eq. 2.19) [18]. The derivation for this equation is presented by Walker in [7].

$$p_r \approx \frac{\rho c^2 Q_0}{V} e^{-j\omega t} \sum_N \frac{\epsilon_{N,x} \epsilon_{N,y} \epsilon_{N,z} \psi(S) \psi(R)}{\frac{2\omega_N \kappa_N}{\omega} + j \left(\frac{\omega_N^2}{\omega} - \omega \right)} \quad (2.19)$$

where:

p_r	=	reverberant field sound pressure (Pa)
ρ	=	air density (kg/m^3)

The total sound pressure (contributions of the reverberant field summed with the direct sound) is calculated using Eqs. 2.20 & 2.21 [7].

$$p_d \approx \frac{\rho}{4\pi r} \left(\sin\left(\frac{\omega r}{c}\right) - j \cos\left(\frac{\omega r}{c}\right) \right) \quad (2.20)$$

$$p_t = p_d + p_r \quad (2.21)$$

where:

p_d	=	direct sound pressure (Pa)
p_t	=	total sound pressure (Pa)

These calculations give theoretical predictions of the pressure distribution across a closed rectangular space. The above equations operate assuming the system consists of a single sound source. When a system consists of multiple sources, superposition can be used to determine the total sound pressure by calculating the contribution from each individual source separately and then summing the contributions.

2.2.3 Spatial variance

The equations in the previous section provide a means for calculating sound pressure at individual locations within a closed rectangular space. It is important to have a method for relating individual locations within a space to determine the overall pressure variation across a listening area. A metric that serves this purpose is referred to as spatial variance [2, 17-19]. Spatial variance is calculated one discrete frequency bin at a time, over multiple listening locations. The frequency bin spatial variances are averaged, giving a single value representing spatial variance over the specified listening locations and frequency bins (Eq. 2.22) [18]. Before calculations are carried out, sound pressure (Pa) must be converted to sound pressure level (dB).

High spatial variance values indicate weak correlation between low-frequency responses across the listening area. This indicates that over a wide-area the listening experience varies greatly, which is the underlying problem concerning small-room low-frequency acoustics.

$$SV = \frac{1}{N_f} \sum_{i=f_{lo}}^{f_{hi}} \sqrt{\frac{1}{N_p - 1} \sum_{p=1}^{N_p} (L_p(p, i) - \overline{L_p(i)})^2} \quad (2.22)$$

where:	SV	=	spatial variance (dB)
	N_f	=	number of frequency bins
	N_p	=	number of listening locations
	f_{lo}, f_{hi}	=	frequency range (Hz)
	$L_p(p, i)$	=	sound pressure level (dB) at listening location, p , and frequency bin, i
	$\overline{L_p(i)}$	=	mean sound pressure level (dB) over all listening locations at frequency bin, i

2.2.4 Magnitude deviation

Another metric used to characterize low-frequency acoustics is magnitude deviation. Magnitude deviation represents how the frequency response at a single location differs from the average response over all listening locations [18, 19]. While spatial variance highlights the extent of spatio-pressure variation, it does not provide detailed insight into which location responses differ the most from the average room response. Magnitude deviation allows for this characterization (Eq. 2.23) and can be averaged over all listening locations to give an alternative metric to spatial variance [18].

$$MD = \sqrt{\frac{1}{N_f - 1} \sum_{i=f_{lo}}^{f_{hi}} (L_p(i) - \overline{L_p})^2} \quad (2.23)$$

where:	MD	=	magnitude deviation (dB)
	$L_p(i)$	=	sound pressure level (dB) at frequency bin, i
	$\overline{L_p}$	=	mean sound pressure level over all frequency bins (dB)

Magnitude deviation is of particular use when evaluating a low-frequency correction method's performance over a large area. Spatial variance gives a single metric indicating the extent of response variation, but magnitude deviation provides more detail by highlighting the areas that contribute most to spatial variance.

2.2.5 Mean output level

Finally, mean output level (MOL) is a metric used to judge system efficiency by comparing pressure output between various configurations. MOL is calculated in decibel (dB) over all frequency bins and listening locations, giving the mean sound pressure level across a listening area (Eq. 2.24) [19].

$$MOL = \frac{1}{N_f N_p} \sum_{i=f_{lo}}^{f_{hi}} \sum_{p=1}^{N_p} L_p(p, i) \quad (2.24)$$

This is especially useful to compare correction systems in terms of efficiency. Generally, it is desirable to have a corrected system MOL close to the uncorrected system MOL. Large differences between uncorrected and corrected system MOLs indicate poor efficiency.

A useful extension to this metric is proposed by Welti in [17] where a correction factor based on the number of sources is applied to the raw MOL value (Eq. 2.25). This is termed the low-frequency factor.

$$LF = MOL - 20\log_{10}(N_s) \quad (2.25)$$

where: LF = low-frequency factor
 N_s = number of active sources

The low-frequency factor provides a dimensionless metric where higher values indicate higher system efficiency. Welti demonstrates that as additional subwoofers are added to a system, the low-frequency factor will decrease due to destructive acoustical interference of the outputs of each individual unit [17]. Any practical multi-subwoofer correction system should be able to maintain a high low-frequency factor, indicating acceptable efficiency.

2.3 Influential factors

Low-frequency room acoustics are greatly affected by factors rooted in the naturally-occurring room acoustics. Some of these factors affect a sound reproduction system's efficiency, specifically MOL and low-frequency factor (presented in the previous section), while others influence the modal frequencies which can be critical when considering the implementation of a low-frequency correction system.

2.3.1 Surface absorption and reactance

The surface properties of the materials in a room strongly influence the acoustical nature of a space. This has been demonstrated to a certain extent by the calculation of the theoretical reverberant field, which is largely due to the total absorption of a room, and also the derivation of the Schroeder frequency formula, where the absorption level has an inverse relationship to the low-frequency upper boundary.

Sound reproduction systems in small rooms rely on the room acoustics to reinforce overall sound output through wall reflections. This is specifically important at low-frequencies where the isolated sound power output from subwoofers may not be adequate. Often the reflective (or reverberant) sound field is stronger than the direct sound, as previously highlighted in Eq. 2.5, showing the role room acoustics can play in the overall perception of reproduced sound.

Given the reverberant field's dominance at low frequencies, a room's absorptive properties strongly shape the reproduced characteristics of a sound. Lack of absorption capabilities in a narrow frequency band can cause an unpleasant resonance in a room, while overly-efficient absorption may cause a room to sound “dead,” or lacking acoustical character.

Unpleasant resonances also arise due to the reactive properties of a surface where certain frequencies excite the surface within a characteristic resonant band, injecting additional energy in that band into the room. This occurs when non-rigid construction materials are used. Rigid materials such as concrete generally do not introduce this variety of resonances.

2.3.2 The Waterhouse effect

An important principle related to low-frequency sound reinforcement is known as the Waterhouse effect, named after Richard Waterhouse, who first described the effect in 1958 [20]. The effect occurs through interaction of direct and reflected sound waves from a source placed in close proximity to a surface. If the source is placed 30 cm from the surface, for example, there will exist a virtual source located 30 cm on the opposite side of the surface. Assuming perfect reflectance (no surface absorption), the reflected wave will interfere with the

direct sound to boost or cut the perceived output at a standard listening distance; an effect which is frequency-dependent based on wavelength of the frequency under inspection.

If a source is placed precisely at the surface, then the source and virtual source exist at the same location. This effectively doubles the pressure output of the source over all frequencies due to the simultaneous wall reflection, corresponding to a 6 dB boost in SPL. Placing a source at a room junction (two surfaces) gives 12 dB SPL boost due to the three virtual sources while room corner placement (three surfaces) gives an 18 dB SPL boost due to the eight virtual sources [2, 20]. The principle of the acoustic center allows a subwoofer to be positioned so that the direct and virtual sources originate at the same location [21, 22].

2.3.3 Helmholtz resonance

This low-frequency room acoustics discussion has assumed closed spaces to this point. Unfortunately, this is not the case in practical situations. Doors and windows (along with ventilation shafts) are regularly opened, eliminating the closed status of the space. Assuming openings are sufficiently small compared to the reproduced wavelengths, a resonance occurs.

This variety of resonance is known as Helmholtz resonance, described as a rigid enclosure of air connected to an external space through a small opening (relative to wavelength) [23]. The coupling of the two spaces results in what is analogous to a mass and spring mechanical system. Equation 2.26 is used to calculate the Helmholtz resonant frequency [23].

$$f_H = \frac{c}{2\pi} \sqrt{\frac{A_H}{V(l_H + \gamma_H)}} \quad (2.26)$$

where:	f_H	=	fundamental Helmholtz resonance (Hz)
	A_H	=	cross-sectional area of the opening (m^2)
	l_H	=	length of the opening (m)
	γ_H	=	$0.8\sqrt{A_H}$ = end correction term

The fundamental resonant frequency found with Eq. 2.26 is only approximate in real rooms, as the geometry of the door, windows and/or the ventilation shafts are complex and not accounted for in the calculation. Nevertheless, Helmholtz resonance has an influence on the acoustical response of a room (especially at very low frequencies) and therefore must be considered when addressing various sound reproduction problems.

2.3.4 Temperature and humidity

Temperature and humidity are influential factors in acoustics that are often ignored. These factors cause variations in the speed of sound in air and also in the level of air absorption. The speed of sound in air due to temperature is calculated using Eq. 2.27 [24].

$$c = \sqrt{\frac{\gamma \bar{R} T}{\bar{m}}} \quad (2.27)$$

where:	T	=	absolute temperature ($^\circ\text{K}$)
	γ	=	1.3987 = ratio of specific heats
	\bar{R}	=	8.31451 J/Kmol = universal gas constant
	\bar{m}	=	0.028904 Kg/mol = mean molar mass, dry air

Most calculations assume a room temperature of 20 °C (293 °K). This gives 343 m/s as the speed of sound. If the temperature were increased to 30 °C (303 °K), then the speed of sound increases to 349 m/s, an increase of 2%. This may not be a major change, but when considering the range of temperatures encountered in outdoor events (0 – 40 °C), the speed of sound may vary up to 7% which causes acoustical calculations based on room temperature to exhibit errors. It is suggested that temperature at a single location within a room can vary by plus or minus 10% due to air flow [24]. Whether this is problematic must be explored when designing room correction procedures with computer simulations to ensure accuracy.

While air humidity is not considered to cause significant changes to the speed of sound (due to minimal change in the mean molar mass [24]), it can cause changes in air absorption, possibly up to 10% [24], which alters the acoustical characteristics of a space. Most closed spaces, however, maintain fairly constant humidity, thus this does not factor greatly into closed space acoustical considerations. As with temperature, humidity plays a significant role for outdoor applications, requiring sound to be transmitted through long distances in the air. Higher humidity levels increase air absorption making it more difficult to transmit sound over long distances.

2.4 Measurement techniques

The various acoustical characteristics mentioned in the previous sections of this chapter require accurate measurement techniques to allow researchers to precisely characterize acoustical spaces. While a number of methods are available in practice, two techniques were chosen for this specific project: maximum length sequences (MLS) and tone bursts. Each of these measurements delivers key data indicating the nature of a room's acoustical response.

2.4.1 Maximum length sequences (MLS)

Maximum length sequences (MLS) are commonly used for acoustic measurements, often when high amplitude background noise (a low signal-to-noise ratio) prevents accurate direct impulse response measurements. The MLS is a pseudo-random binary signal that consists of a repeatable pattern, generated using a feedback shift register with values determined using recursion relations [25]. The MLS exhibits a flat frequency spectrum up to the Nyquist frequency (half the sampling rate), excluding the DC term.

MLS generation begins by defining a recursion relation. A primitive polynomial for an MLS signal of the 4th order is shown in Eq. 2.28 [25].

$$h(x) = x^4 + x + 1 \quad (2.28)$$

The feedback shift register for 4th order MLS generation is shown in Fig. 2.3, where a_i and a_{i+1} correspond to x and 1 in Eq. 2.28, respectively. The x^4 term does not play a role in the summing procedure, as the process has only four elements, hence the 4th order descriptor.

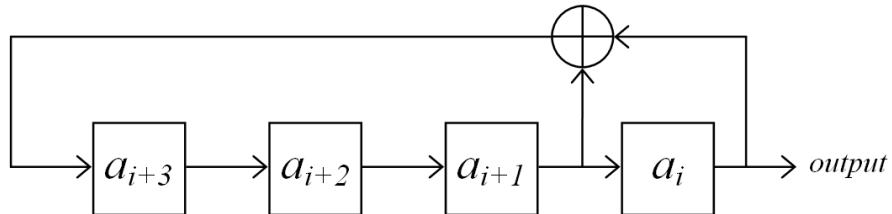


Fig. 2.3 Feedback shift register for 4th order MLS generation [25]

Based on the shift register structure shown in Fig. 2.3 and the primitive polynomial in Eq. 2.28, the recursion relation can be written as:

$$a_{i+4} = a_{i+1} + a_i \quad (2.29)$$

where the term a_{i+4} is calculated using a_{i+1} and a_i and held in temporary memory until the registers are shifted to the right, giving an output of the value previously in a_i and moving a_{i+4} to the a_{i+3} element. Since an MLS is binary by definition, the output values of the shift register must equal zero or one. To avoid non-binary values, a_i and a_{i+1} are run through an exclusive OR (XOR) gate, as opposed to using a simple summation.

The initial state of the register elements is unimportant since the shift register cycles through every non-zero four-digit binary combination over the entire MLS signal and is periodic. However, an initialization of all zeros cannot be used since the XOR gate would never generate a one [25]. For illustrative purposes, the elements were initialized to all ones and the first ten MLS samples were generated with the system in Fig. 2.3 and displayed in Table 2.1 [25]. An MLS signal will be $2^m - 1$ samples long, where m is the MLS order.

Time step (i)	Elements ($a_{i+3}, a_{i+2}, a_{i+1}, a_i$)	MLS value
0	1111	-
1	0111	1
2	0011	1
3	0001	1
4	1000	1
5	0100	0
6	0010	0
7	1001	0
8	1100	1
9	0110	0
10	1011	0

Table 2.1 First ten output values of a 4th order MLS

When taking measurements using a loudspeaker it is ideal to have a signal that covers a transducer's full operating range (i.e. signal amplitude of -1 to 1). The MLS is, by definition, in the range of 0 to 1. To rectify this, the MLS can be converted to a -1 to 1 scale by setting all zeros to 1 and all ones to -1. This is accomplished using Eq. 2.30 [25].

$$s_i = (-1)^{MLS_i} \quad (2.30)$$

where: s_i = adjusted MLS element at time step, i
 MLS_i = raw MLS element at time step, i

Impulse response extraction is relatively straightforward for an MLS signal. The measured signal is by definition the convolution of the source signal and the room impulse response (Eq. 2.31).

$$y = s * h \quad (2.31)$$

where y is the measured signal, with s and h representing the MLS source signal and room impulse response, respectively. Since the goal is to determine the room impulse response, h , correlation with respect to the MLS source signal, s , can be applied to Eq. 2.31 (Eq. 2.32).

$$\phi_{sy} = \phi_{ss} * h \quad (2.32)$$

where: ϕ_{sy} = correlation of the MLS and measurement
 ϕ_{ss} = auto-correlation of the MLS signal

Since correlation of the MLS with itself (auto-correlation) gives a perfect impulse (flat frequency response), the equation can be reduced to indicate that the impulse response of the system is equal to the correlation between the input (MLS) and output (measured) signals (Eq. 2.33) [25].

$$h = \phi_{sy} \quad (2.33)$$

Measurements using only one variation of an MLS signal are not always sufficient. Measurements contain harmonic distortion which can reduce measurement accuracy. This problem is diverted by generating multiple MLS signals of the same order, but with slightly different recursion relations [26]. The multiple room impulse responses extracted from the MLS measurements are averaged to suppress the unwanted distortion.

An effective method to accomplish this discards the maximum and minimum values of the measurements at each time step and averages over the remaining data. The effectiveness of this technique is highlighted in Fig. 2.4 [27], where the bold blue line represents the mean response which does not exhibit any sharp peaks due to distortion (peaks which are present in the individual measurements in Fig. 2.4).

With harmonic distortion suppressed, the impulse response of the system can be used to determine the frequency response. This technique is very effective for analyzing the steady-state characteristics of a system, assuming that a sufficiently long MLS is utilized, so that the room has time to reach steady-state [25].

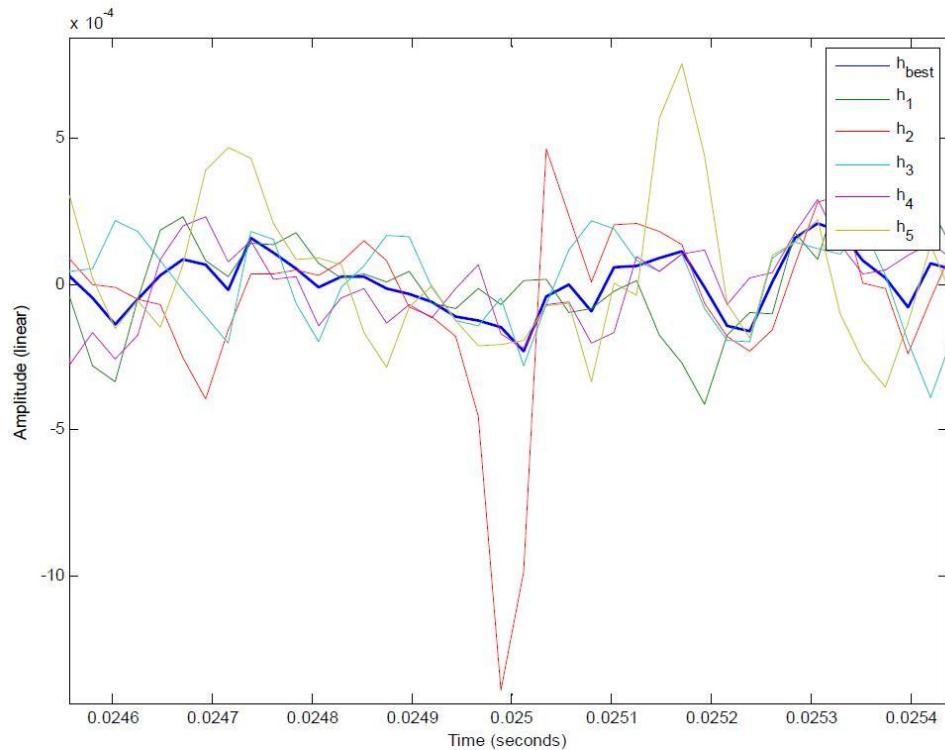


Fig. 2.4 Magnification of the extracted impulse responses from five separate MLS measurements and a smoothed response after averaging (solid blue curve) [27]

2.4.2 Tone bursts

Tone bursts, while less popular than MLS, are used specifically for transient measurements of a system. This technique has been used for many years, where one of the first applications utilized tone bursts to determine the correlation of transient measurements of this type to subjective test data [28], while a decade later Kaminsky used tone bursts to measure the transient behavior of loudspeakers [29].

Corrington and Kaminsky's applications only considered transient characteristics, however Linkwitz conducted significant work applying tone bursts to determine transient and steady-state characteristics [30-32]. The initial goal of these tests was to evaluate loudspeakers by removing the room effects on the response, giving purely anechoic data [30].

Tone bursts are simple signals to generate. A pure sinusoid is generated over a specified number of full cycles. Next, a window is applied through multiplication. Any variety of window can be used. Linkwitz employs a raised cosine window in his work (Eq. 2.34) [31].

$$w(t) = \frac{1}{2} - \frac{1}{2} \cos\left(2\pi \frac{t}{T_w}\right) \quad (2.34)$$

where: $w(t)$ = raised cosine window
 T_w = time length of the window (s)

The burst is repeated a given number of times (usually around 3 – 5) and simultaneously measured at one or more listening location. For multi-burst signals the modulation rate is a fixed percentage of the carrier frequency (Eq. 2.35) [31].

$$f_m = f_{tb}/N_c \quad (2.35)$$

where: f_m = tone burst modulation frequency (Hz)
 f_{tb} = tone burst frequency (Hz)
 N_c = number of sinusoidal cycles per burst

The number of cycles in each tone burst repetition determines two things. First, it determines the frequency resolution of the test signal. For example, a ten-cycle burst with an envelope generated using Eq. 2.34 covers one-third an octave with a center frequency at the burst frequency. A wider burst shape gives higher frequency resolution, but this is not necessary or desirable when focusing on transient characteristics [32].

The second aspect of the test signal controlled by the number of cycles (in conjunction with window shape) dictates whether the test signal measures transient or steady-state characteristics. A short burst is ideal for transients, while a longer burst allows a room to reach steady-state. According to Linkwitz, the ideal number of cycles for steady-state measurements in a small room is around forty cycles, while transient measurements are best taken utilizing five- to fifteen-cycle bursts [32].

An interesting observation by Linkwitz is that certain source/receiver position combinations cause the perception of twice the original modulation frequency, but still have the perception of the correct burst frequency (Fig. 2.5). Other positions give accurate perception of the modulation frequency, but with a slightly inaccurate perceived burst frequency [32].

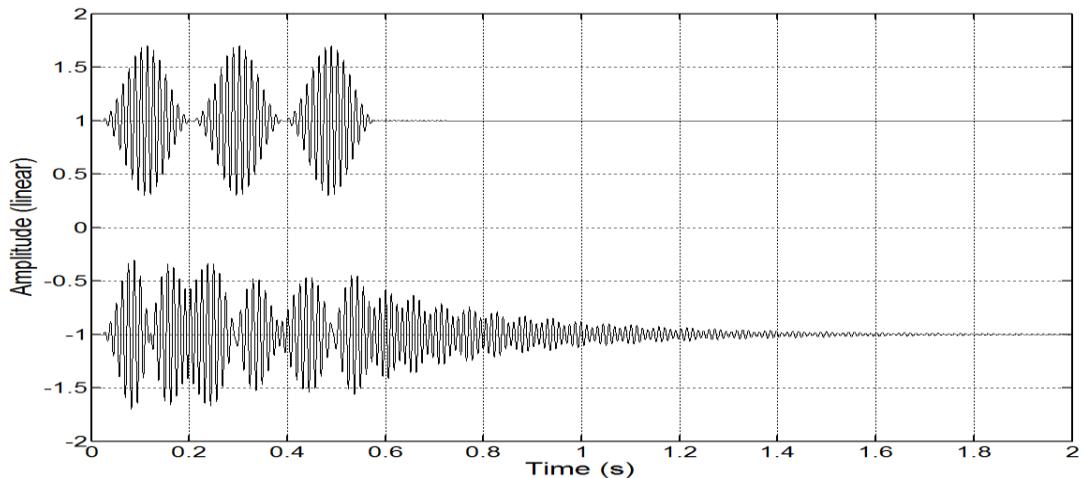


Fig. 2.5 Tone burst input-output comparison showing modulation frequency doubling
(top = input tone burst, bottom = measured response at the listening location)

These observations highlight the advantages of tone bursts, as certain transient characteristics may not be noticed with other forms of test signals, such as MLS. Linkwitz suggests that using test signals that only measure steady-state characteristics may not indicate a noticeable difference between omnidirectional and dipole sources. The major differences often lie in transient characteristics [32].

It is clear that tone-burst testing is necessary to determine the advantages/disadvantages between sources of varying polar patterns. Transient characteristics are likely critical to these differences, while the steady-state may not be drastically affected.

2.5 Chapter summary

There are numerous aspects influential to room acoustics, specifically in the low-frequency band. Room dimensions and construction material properties define the low- and high-frequency band ranges. In the low-frequency band, these factors contribute to the occurrence of standing waves at certain frequencies, known as room-modes. Room-modes result in spatiotemporal variation in the frequency response, which complicates matters when attempting to deliver equal listening experiences to all individuals scattered over a wide area.

Two useful measurements used to analyze a room's acoustic properties are the MLS and the tone burst. MLS measurements give accurate steady-state impulse response data for points in a room, allowing for accurate frequency response inspection. The MLS also has the advantage of being insensitive to background noise, as it can function with very low signal-to-noise ratios. Tone bursts are useful for inspecting transient behavior at individual frequencies. Tone burst measurements highlight waveform degradation that occurs between the source and listener location. By altering the number of cycles per burst, these measurements can be used for steady-state characteristics as well, but are less useful than MLS since they can only be performed at one frequency at a time.

The discussed issues in small-room acoustics must be considered when designing low-frequency room response correction systems. Failure to address these aspects can result in ineffectiveness, inefficiency or instability. The topics discussed in this section are addressed in forthcoming chapters as existing room-mode correction methods are analyzed and novel methods developed.

3 Low-frequency modeling

Software that accurately simulates room acoustics can be a powerful research tool, saving time, materials and funding that would otherwise be required to implement different systems in the real world. A great deal of research has been carried out on acoustics simulation methods over the past half-century, resulting in a wide-variety of techniques. Each of these techniques offers its own advantages and disadvantages, which must be taken into account when choosing an appropriate approach. This chapter first highlights the most common acoustical simulation methods, followed by a detailed description of the chosen simulation technique for this work and concluded with a description of an original simulation toolbox developed in MATLAB.

3.1 Simulation methods

Acoustical simulation methods can be divided into three general categories: geometrical, statistical and hybrid. Geometrical approaches consider the physical layout of a space to determine how sound waves interact with various surfaces in terms of absorption and reflections. Statistical approaches utilize a discretized form of the wave equation to model sound wave propagation throughout a space, commonly making geometrical approximations, but providing more detailed wave propagation data. Hybrid methods combine geometrical and statistical approaches. These techniques are configured to apply statistical modeling for the low-frequency range for detailed wave behavior which is important in analyzing modal behavior within a space, while the high-frequency range is simulated with a geometrical approach to ensure there are no negative effects due to any geometrical approximations (especially with curved surfaces). Geometrical approaches are usually computationally-efficient at higher frequencies, so hybrid methods run faster than a full-range statistical approach. The following sections highlight modeling techniques used in acoustics research.

3.1.1 Image source

Image source acoustical modeling is of the geometrical variety, operating on the assumption that surface reflections can be modeled by reconstructing a space using virtual sources that mirror the primary source along the plane of the reflective surface [33]. This method gained popularity within the acoustics community in the early 1970s due to its simplicity and correspondingly high computational efficiency [34, 35].

The simplest form of image source modeling treats room boundaries as perfectly rigid, allowing a surface reflection to be represented by an image placed symmetrically on the opposite side of said surface. This can be expanded to higher order reflections where image sources from lower-order reflections are included. The approximate number of image sources is calculated based on the required simulation radius, defined by the simulation time (Eq. 3.1) [33]. This principle is represented graphically in Fig. 3.1 (adapted from [35]). The images within the dotted circle are included in the model.

$$N_r = \frac{4\pi c^3}{3V} T^3 \quad (3.1)$$

where: N_r = number of reflections (image sources)
 T = simulation time (s)

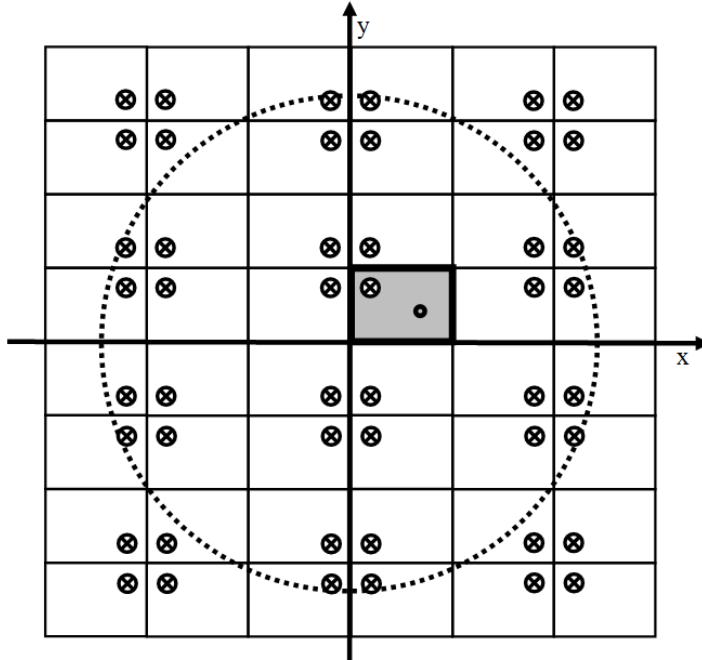


Fig. 3.1 Image source spatial configuration
(circle = receiver, crosses = sources, gray box = modeled room)

Once the image source positions are known, the propagation delay can be calculated relative to the receiver location(s). The received signal is constructed from these values via a summation of delayed versions of the source signal. If the reflection coefficients of the walls are not equal to one (perfectly reflective) then the original source signal must be attenuated based on how many walls each image source traverses through before arriving at the receiver.

While this technique is useful for generating impulse responses for single listener locations, it can prove to be inefficient when multiple locations are required, as each location requires multiple summations of the source signal, increasing memory needs. Also, if a room is nonrectangular, the number of images increases rapidly, as demonstrated by Eq. 3.2 [33].

$$N_r = 1 + \frac{n_s}{n_s - 2} ((n_s - 1)^{m_r} - 1) \approx (n_s - 1)^{m_r} \quad (3.2)$$

where:

n_s	=	number of surfaces
m_r	=	reflection order

Due to this complexity, image source models are generally used for rectangular geometries where a solution based on low-order reflections suffices [33].

3.1.2 Ray tracing

Like the image source method, ray tracing is of the geometrical variety of acoustical modeling and operates by calculating reflections due to the surfaces of a space. While image source modeling achieves this by placing a number of virtual sources at locations outside the simulation space to approximate the propagation path of a reflected sound wave, ray tracing operates by scattering sound particles in various directions from the source which are traced throughout the room as they bounce off the surfaces. As with image sources, when each particle encounters a surface it loses energy based on the absorption coefficient of the surface and its

reflection angle is determined by the angle of incidence [33]. This method was first described by Krokstad et al in 1968 [36].

The response at a receiver is determined by defining a small area surrounding the receiver that serves to catch the reflected (and direct) sound particles. Since a method like this runs the risk of not catching all reflections in this manner, it is necessary to define a minimum number of elements (or rays) that must be emitted from each source to generate an accurate response at the receiver (Eq. 3.3) [33].

$$N_{rays} \geq \frac{8\pi c^2}{A_R} t^2 \quad (3.3)$$

where:	N_{rays}	=	minimum number of rays required
	t	=	analysis time (s)
	A_R	=	surface area of the receiver (m^2)

A problem with ray tracing is choosing an appropriate ray shape. The goal is to find a shape (with the accompanying density function) that provides maximal accuracy. At the moment, triangular pyramids are the common ray shape, but this is likely to change as alternate approaches are developed [33].

In terms of low-frequency acoustical modeling, ray tracing encounters issues because the propagation distance of the majority of the reflections is less than the wavelengths in the low-frequency band, resulting in an incomplete representation of the overall frequency response. The method does benefit, however, from good accuracy in the high-frequency band and is computationally efficient.

3.1.3 Finite element method (FEM) and boundary element method (BEM)

One of the more popular statistical simulation methods is the finite element method (FEM). FEM is used extensively in structural engineering applications where precise mathematical calculations are difficult due to complex geometries and material properties. This method is used to approximate these difficult problems by splitting the area of interest (acoustical space, physical structure, etc) into small interlocking elements. These elements are connected by nodes which can be points (1D), lines (2D) or surfaces (3D) [37]. Once the elements and nodes have been initialized (a process called meshing), equations are defined to determine the value of interest for each element. Generally, neighboring elements influence each other based on what are called shape equations. With the shape and property equations known, an overall solution can be determined by summing all individual elements [37].

FEM accuracy is limited by the size of the elements. Smaller elements allow for accuracy at higher frequencies, but require greater computational resources. For low-frequency acoustical analysis, however, this should not be an issue as wavelengths are sufficiently large to allow for a small number of elements while maintaining reasonable accuracy. FEM operates in the frequency domain and it has been suggested that it can be difficult to accurately convert FEM frequency domain results to the time domain [38].

In cases where only the response at surfaces needs to be modeled, the boundary element method (BEM) is more efficient than FEM. BEM only considers the boundary elements in a model while operating using the same principles as FEM. This allows for three-dimensional FEM models to be scaled to two-dimensional problems, reducing the computational complexity. This method is practical for loudspeaker horn design and has been used extensively for such work in

the past [39]. As with FEM, this method operates in the frequency domain and can prove difficult to convert to the time domain.

3.1.4 Hybrid methods

Given the shortcomings of each of the described acoustical modeling methods, it is common practice to combine methods to form a hybrid system. Generally, these systems consist of statistical modeling for the low-frequency range (up to the Schroeder frequency) and geometrical modeling for the remainder of the bandwidth [33]. This avoids the low-frequency problems encountered with geometrical methods and also limits computational complexity of the statistical methods since they are only required to operate in the low-frequency band, necessitating relatively few elements.

Hybrid systems consisting of only statistical methods have also been developed. These systems combine FEM and BEM where the FEM portion of the model simulates the non-boundary conditions of the acoustical space and the BEM portion concentrates only on the boundaries. This allows for the BEM elements, which only exist in a two-dimensional space to be smaller, giving finer detail of surface diffraction and scattering while the FEM elements remain larger to maintain system efficiency [40].

While hybrid models are valuable for full-range simulations, they are not necessary for low-frequency analysis. For these cases, a well-chosen statistical method will suffice.

3.2 Finite-difference time-domain theory

The finite-difference time-domain (FDTD) simulation method has existed since the 1960's in electromagnetics research, but has only begun to gain popularity in acoustics over the past two decades, with early work conducted by Botteldooren in 1994 [41]. FDTD's appeal concerning acoustical modeling is that it operates in the time domain, allowing for transient analysis and simple conversion to the frequency domain [38]. Also, the method is relatively straightforward to implement, unlike FEM or BEM, functioning on a small set of equations that can be expanded to include the desired system complexity.

3.2.1 Structure

FDTD operates over a set of offset grids. For a two-dimensional simulation, this set consists of a single sound pressure grid and two particle velocity grids (one for the x-dimension and one for the y-dimension). The elements of these grids are uniformly spaced over each dimension (not necessarily the same spacing in each dimension, though) with the particle velocity grids offset from the sound pressure grid by half of the spacing interval (Fig. 3.2).

In addition to the spatial offset of the grids, the elements are also offset temporally. The simulation begins with a predefined sound pressure time sequence at the source location(s). These points are used to calculate the adjacent particle velocity values in between the pressure grid time steps. The calculated particle velocity elements are then used in turn to update the pressure grid on the time step. This computation process continues until the simulation has reached its conclusion. Since the grid points are calculated recursively, only the two most recent sets of elements must be stored in memory; the others may be discarded, unless they are required for post-processing.

The FDTD update equations operate on a discretization of the wave equation using coupled first-order differential equations by computing the derivative with the central finite difference

[38]. This means that a straight line is placed between values at neighboring points (in time or space) and the slope is taken as the estimate of the first derivative.

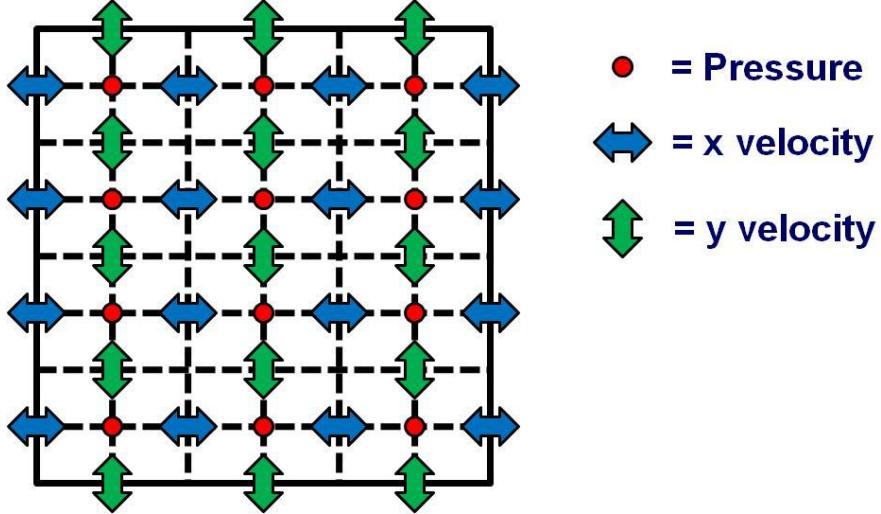


Fig. 3.2 Grid layout of a 2D FDTD simulation

Non-boundary elements are updated using a set of four equations for a three-dimensional simulation, using zeroed pressure and particle velocity grids at $t = 0$ as the initial condition (Eqs. 3.4 – 3.7, expanded from 2D equations in [42]). These equations are valid for all non-boundary elements. Boundary condition handling is discussed in the following section.

$$u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) = u_{x+\frac{\Delta x}{2},y,z}^x \left(t - \frac{\Delta t}{2} \right) - \frac{\Delta t}{\rho \Delta x} [p_{x+\Delta x,y,z}(t) - p_{x,y,z}(t)] \quad (3.4)$$

$$u_{x,y+\frac{\Delta y}{2},z}^y \left(t + \frac{\Delta t}{2} \right) = u_{x,y+\frac{\Delta y}{2},z}^y \left(t - \frac{\Delta t}{2} \right) - \frac{\Delta t}{\rho \Delta y} [p_{x,y+\Delta y,z}(t) - p_{x,y,z}(t)] \quad (3.5)$$

$$u_{x,y,z+\frac{\Delta z}{2}}^z \left(t + \frac{\Delta t}{2} \right) = u_{x,y,z+\frac{\Delta z}{2}}^z \left(t - \frac{\Delta t}{2} \right) - \frac{\Delta t}{\rho \Delta z} [p_{x,y,z+\Delta z}(t) - p_{x,y,z}(t)] \quad (3.6)$$

$$\begin{aligned} p_{x,y,z}(t + \Delta t) = & p_{x,y,z}(t) - \frac{c^2 \rho \Delta t}{\Delta x} \left[u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) - u_{x-\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) \right] \\ & - \frac{c^2 \rho \Delta t}{\Delta y} \left[u_{x,y+\frac{\Delta y}{2},z}^y \left(t + \frac{\Delta t}{2} \right) - u_{x,y-\frac{\Delta y}{2},z}^y \left(t + \frac{\Delta t}{2} \right) \right] \\ & - \frac{c^2 \rho \Delta t}{\Delta z} \left[u_{x,y,z+\frac{\Delta z}{2}}^z \left(t + \frac{\Delta t}{2} \right) - u_{x,y,z-\frac{\Delta z}{2}}^z \left(t + \frac{\Delta t}{2} \right) \right] \end{aligned} \quad (3.7)$$

where:

u^x	=	x-dimension particle velocity (m/s)
u^y	=	y-dimension particle velocity (m/s)
u^z	=	z-dimension particle velocity (m/s)
$p_{x,y,z}$	=	sound pressure at point (x, y, z) (Pa)
t	=	time (s)
Δt	=	time step (s)
Δx	=	element spacing in the x-direction (m)
Δy	=	element spacing in the y-direction (m)
Δz	=	element spacing in the z-direction (m)

3.2.2 Boundary conditions

Upon inspection of the particle velocity equations (Eqs. 3.4 – 3.6), it is clear that it is impossible to calculate particle velocity values at boundaries because there is only one of the two required adjacent pressure values available. In order to properly handle boundary conditions, the model must utilize an asymmetric finite difference approximation for the spatial derivative [38, 42]. This approximation is expressed by Eq. 3.8 (x-dimension example shown). The last term in Eq. 3.8 is the unknown pressure value and must be expressed as the product of the particle velocity at the element near the wall (perpendicular velocity component) and the characteristic wall impedance (Eq. 3.9) [42]. Since complex wall impedance data is rarely available, the characteristic wall impedance is assumed to be real, and it has been argued that the real component is the dominant factor [41]. The characteristic wall impedance, Z , is calculated using Eq. 3.10 [41].

$$u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) = u_{x+\frac{\Delta x}{2},y,z}^x \left(t - \frac{\Delta t}{2} \right) - \frac{\Delta t}{\rho \Delta x} [p_{x-\Delta x,y,z}(t) - p_{x,y,z}(t)] \quad (3.8)$$

$$u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) = u_{x+\frac{\Delta x}{2},y,z}^x \left(t - \frac{\Delta t}{2} \right) - \frac{\Delta t}{\rho \Delta x} \left[p_{x-\Delta x,y,z}(t) - Z u_{x+\frac{\Delta x}{2},y,z}^x(t) \right] \quad (3.9)$$

$$Z = \rho c \frac{1 + \sqrt{1 - \alpha}}{1 - \sqrt{1 - \alpha}} \quad (3.10)$$

Equation 3.10 shows that the characteristic wall impedance, Z , approaches infinity when the absorption coefficient, α , is very small. Olesen handles with this by assuming that the particle velocity changes linearly over adjacent time steps (Eq. 3.11). Substituting this assumption into Eq. 3.9 gives the final boundary condition for a rigid surface that is stable (Eq. 3.12), due to the elimination of the dependence on the absorption-sensitive wall impedance equation (Eq. 3.10) [42].

$$u_{x+\frac{\Delta x}{2},y,z}^x(t) = \frac{u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) + u_{x+\frac{\Delta x}{2},y,z}^x \left(t - \frac{\Delta t}{2} \right)}{2} \quad (3.11)$$

$$\begin{aligned} u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) \\ = \frac{\frac{\rho \Delta x}{0.5 \Delta t} - Z}{\frac{\rho \Delta x}{0.5 \Delta t} + Z} u_{x+\frac{\Delta x}{2},y,z}^x \left(t - \frac{\Delta t}{2} \right) + \frac{2}{\frac{\rho \Delta x}{0.5 \Delta t} + Z} p_{x-\Delta x,y,z}(t - \Delta t) \end{aligned} \quad (3.12)$$

For simplicity, the fraction containing the air density term, ρ , is lumped into a constant (Eq. 3.13). This simplifies the positive x-dimension boundary condition to Eq. 3.14. The negative x-boundary condition equation is achieved by reversing the direction of the particle velocity component calculated using the pressure approximation (Eq. 3.15).

$$R_x = \frac{\rho \Delta x}{0.5 \Delta t} \quad (3.13)$$

$$u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) = \frac{R_x - Z}{R_x + Z} u_{x+\frac{\Delta x}{2},y,z}^x \left(t - \frac{\Delta t}{2} \right) + \frac{2}{R_x + Z} p_{x-\Delta x,y,z}(t - \Delta t) \quad (3.14)$$

$$u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) = \frac{R_x - Z}{R_x + Z} u_{x+\frac{\Delta x}{2},y,z}^x \left(t - \frac{\Delta t}{2} \right) - \frac{2}{R_x + Z} p_{x-\Delta x,y,z}(t - \Delta t) \quad (3.15)$$

Similar calculations are used for the y- and z-dimensions. A complete set of equations for the FDTD simulation is contained in Appendix A.

3.2.3 Advanced boundary conditions

The described boundary handling assumes surfaces to be perfectly rigid. In other words, surfaces exhibit frequency-independent properties. Advanced versions of these boundary conditions have been explored in the literature allowing for frequency-dependent absorptive properties and also reactive (non-rigid) surfaces [41, 44, 45].

The most common implementation of advanced boundary conditions converts each boundary element to act as low-order infinite impulse response (IIR) filters [44]. This requires time-dependence where previous pressure values at the boundary (and particle velocity values adjacent to the boundary) are stored and used within the IIR filter structure to calculate reflections. This, of course, requires additional computational memory, decreasing model efficiency.

Alternatively, advanced boundary conditions can be implemented using a classical electromagnetics approach for signal transmission through two or more mediums with different characteristic impedances [45]. This model does not terminate at the surface elements, but carries on to model the entire structure of the surface and could possibly expand to model adjacent rooms. As in electromagnetics, the reflected and transmitted waves are calculated based on the material properties, including the angle of reflectance and transmission. The characteristic impedances can be complex, allowing for frequency-dependent characteristics. This implementation, as with other advanced methods, provides more precise simulation results at the cost of more computational power.

As this phase of research aims to use modeling only to achieve approximate results that can be applied to shape practical solutions to low-frequency control problems, perfectly rigid surfaces should suffice. Validation of the FDTD technique is presented in Section 3.4.

3.2.4 Stability

Care must be taken when defining the element spacing and time step to avoid spectral and/or spatial aliasing. A smaller time step allows for modeling up to higher frequencies, governed by the Nyquist frequency, which requires at least two sample points per frequency wavelength to accurately reproduce the signal. Less than two sample points per wavelength introduces aliasing where the frequency is folded back over the Nyquist frequency causing incorrect identification of a component at a lower frequency [43]. Additionally, the element spacing must be sufficiently less than the smallest wavelength within the bandwidth of interest. It is suggested that reasonable results can be achieved with five to ten elements per wavelength [38]. An element spacing of 10 cm, for example, produces accurate results up to 600 Hz, by this measure.

The time step is calculated based on the user-defined element spacing to ensure stability. It is important to choose a time step, Δt , that corresponds to the element spacing to allow for accurate sound wave propagation and to minimize grid dispersion errors. This is calculated using the following equation [38].

$$\Delta t \leq 1/c \sqrt{\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2}} \quad (3.16)$$

Since Eq. 3.16 gives the maximum time step allowed for system stability, it can be used to ensure stability and maximal computational efficiency.

3.2.5 Non-rectangular FDTD simulation

As few spaces in the real world are perfectly rectangular, it is necessary to provide a modeling method capable of handling non-rectangular topologies. The simplest of the three most common methods ignores elements outside a defined boundary. In the case of a curved boundary, this results in a step-like spatial discretization (Fig. 3.3).

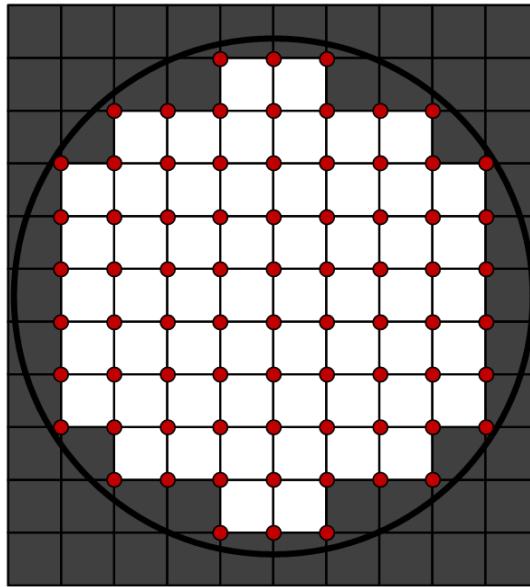


Fig. 3.3 Discretization of an element grid for a circular 2D space (red dots = pressure elements)

The approximations made in this process can cause inaccuracies in terms of acoustic scattering at the boundaries. These inaccuracies are largely frequency-dependant, and worsen with increasing frequency [46].

An advanced version of the above spatial discretization employs a locally-conformal grid to reduce scattering errors. This technique, proposed in [46], allows for variable element spacing, which is used to provide a denser concentration of elements near room boundaries, where scattering accuracy is crucial, while keeping element density as low as possible for non-boundary conditions, increasing computational efficiency (Fig. 3.4).

Locally-conformal element spacing operates by inspecting if the space between two potential elements is intersected by a boundary. If yes, an additional element is placed at the midpoint between the two original elements. This process is repeated over the element-to-element interval until the minimum element spacing, as defined by the user, is reached. The process is performed on all adjacent elements within the grid structure [46]. This requires care when computing pressure and particle velocity updates as the element spacing values (Δx , Δy and Δz) are not constant. The increased resolution around curved room boundaries improves accuracy, especially concerning scattering at high frequencies, as demonstrated in [46]. This benefit is similar to employing an FEM/BEM hybrid model.

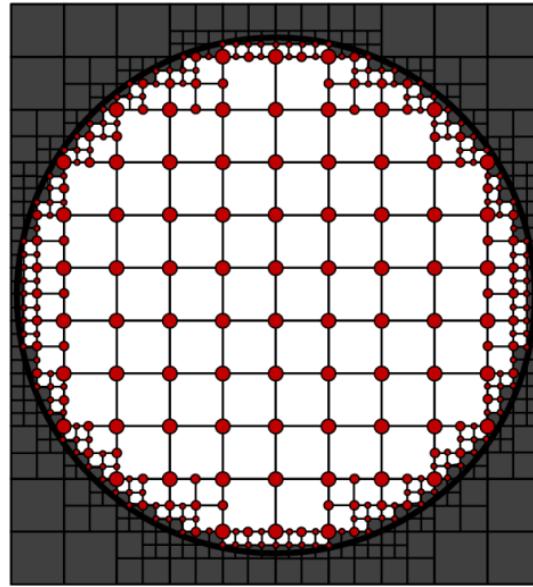


Fig. 3.4 Locally-conformal discretization of an element grid for a circular 2D space
(red dots = pressure elements)

A third approach to non-rectangular topology handling has existed since the early research concerning FDTD applications in acoustics [41]. The technique takes a rigid Cartesian grid structure of elements set up according to the approach detailed in Fig. 3.3 and shifts elements adjacent to boundaries so that they are closer, essentially rounding the grid structure around non-rectangular edges and surfaces (Fig. 3.5).

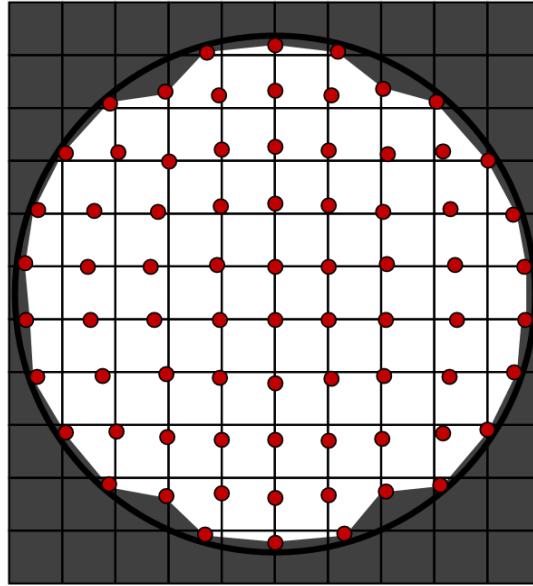


Fig. 3.5 Quasi-Cartesian discretization of an element grid for a circular 2D space
(red dots = pressure elements)

This technique was developed largely due to the low computing capabilities of the day where finer element spacing would be too computationally demanding. The quasi-Cartesian grid layout was shown to give more accurate spectral and temporal results for grids of widely spaced elements as compared to Cartesian layouts [41]. Again, care must be taken with this approach to ensure the spacing variables in the equations are correct for each element pair/trio.

As this project is concerned primarily with low-frequency analysis (below ~ 200 Hz) and is not focused on developing a broadband modeling routine, the Cartesian grid layout was selected as it provides acceptable low-frequency accuracy for modal analysis [38, 42].

3.2.6 Non-rectangular and room obstacle grid masking

Although non-rectangular topology modeling is presented in existing research [41, 46] there is no detailed published explanation regarding the specifics of element grid generation. It is vital to ensure obstacles within a space are handled properly regarding boundary element designation, which has not been discussed in any of the reviewed literature.

In order to effectively implement non-rectangular topology and obstacle handling into the existing FDTD modeling routine, element grid masking was chosen (which is well-suited for MATLAB matrix operations). All examples in the remainder of this section are presented for two-dimensional scenarios with the understanding that the technique can be expanded to three dimensions.

The first step in mask generation is to define the pressure element grid (Fig. 3.6). This is entirely user-controlled and does not involve any special algorithms. For illustrative purposes, the individual grid elements are represented by boxes, rather than dots, to enforce the matrix implementation within MATLAB.

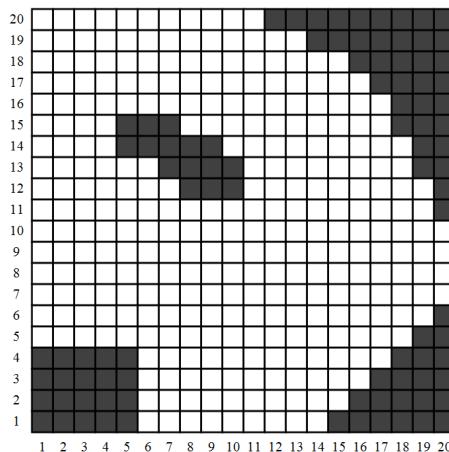


Fig. 3.6 Pressure element grid mask example
(white boxes = inside space (1), gray boxes = outside space (0))

The gray boxes in Fig. 3.6 represent pressure elements lying outside the non-rectangular acoustical space and are assigned a zero value in the mask matrix. All white boxes are pressure elements within the acoustical space and are assigned a one in the mask matrix. All pressure elements are non-boundary, by definition, so no additional masks need to be generated in terms of pressure elements for the pressure update equation (Eq. 3.7).

Separate x- and y- particle velocity element grid masks are generated from the user-defined pressure element grid mask since the grids each have different dimensions, as illustrated in Fig. 3.2. A pressure element grid of dimensions 20x20 (x by y) corresponds to an x-particle velocity element grid of dimensions 21x20 and a y-particle velocity element grid of dimensions 20x21. This is due to the fact that pressure elements are always surrounded by particle velocity elements, requiring an additional element in each primary dimension. With this in mind, particle velocity element grid masks are generated from the pressure element grid mask by searching the pressure element mask for any non-zero values. The corresponding indices of

these non-zero values are then addressed in the particle velocity mask matrices (initialized to all ones) so that these indices are set to zero (outside the non-boundary element set) along with an additional zeroed element adjacent to the final pressure mask defined zeroed element in the positive moving direction (Fig. 3.7).

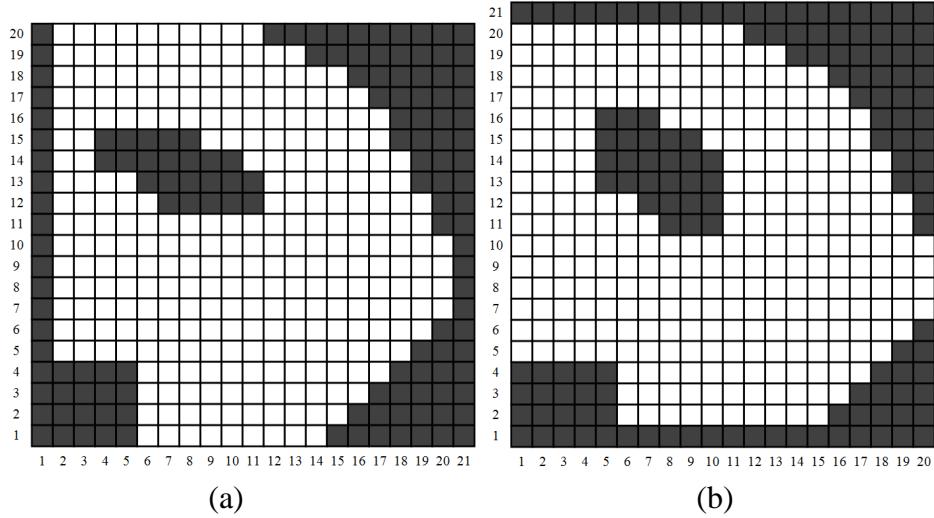


Fig. 3.7 Non-boundary (a) x- and (b) y-dimension particle velocity grid masks
(white boxes = inside space (1), gray boxes = outside space (0))

Next, the non-boundary particle velocity masks are used to generate boundary masks which indicate the particle velocity elements that must be updated with the boundary element equations. These masks are generated by inspecting the non-boundary masks and setting any elements marked outside the space, but adjacent to an element within the space, to be inside the boundary space (set to 1). All other elements are set to zero (outside the boundary space). These masks are valid for use within all particle velocity boundary condition update equations (Fig. 3.8, only x-dimension examples shown).

Additionally, masks must be generated for the non-boundary and boundary particle velocity grids for use within the pressure update equation, which does not operate using the same dimensions as the particle velocity update equations. An additional pressure grid mask must be generated for use within the particle velocity update equations. These masks can be generated by addition or removal of a row/column of non-inclusive elements to meet the dimension requirements. A full set of masks for the two-dimensional example is found in Appendix A.

Using this matrix implementation of FDTD modeling, the individual element update equations (Eqs. 3.4-3.7, 3.14-3.15) can be simplified to operate over all elements in one pass using the masks (Eqs. 3.17 – 3.23, two-dimensional example). Note that the ‘ \circ ’ operator represents an entry-wise product of two matrices of equal dimensions (known as a Hadamard product). This operation is used to represent the masking of the data matrices. Additionally, it is assumed that these equations are for use within a programming environment whereby masked matrices (such as $u_x \circ u_{xM}$) are updated using their previously stored values (also represented by $u_x \circ u_{xM}$ with this example).

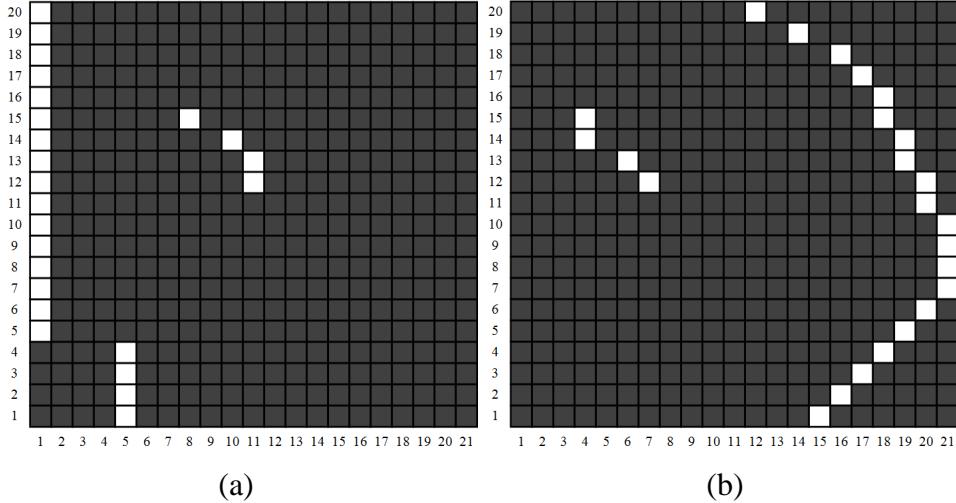


Fig. 3.8 X-dimension boundary condition particle velocity masks for
 (a) negative x-direction direction and (b) positive x-direction
 (white boxes = inside space (1), gray boxes = outside space (0))

- Non-boundary particle velocity update equations:

$$u_x \circ u_{xM} = u_x \circ u_{xM} - \frac{\Delta t}{\rho \Delta x} (p \circ p_{Mx1} - p \circ p_{Mx2}) \quad (3.17)$$

$$u_y \circ u_{yM} = u_y \circ u_{yM} - \frac{\Delta t}{\rho \Delta y} (p \circ p_{My1} - p \circ p_{My2}) \quad (3.18)$$

- Negative (left) x-direction particle velocity update equation:

$$u_x \circ u_{ML} = \frac{R_x - Z_L}{R_x + Z_L} (u_x \circ u_{ML}) - \frac{2}{R_x + Z_L} (p \circ p_{ML}) \quad (3.19)$$

- Positive (right) x-direction particle velocity update equation:

$$u_x \circ u_{MR} = \frac{R_x - Z_R}{R_x + Z_R} (u_x \circ u_{MR}) + \frac{2}{R_x + Z_R} (p \circ p_{MR}) \quad (3.20)$$

- Negative (bottom) y-direction particle velocity update equation:

$$u_y \circ u_{MB} = \frac{R_y - Z_B}{R_y + Z_B} (u_y \circ u_{MB}) - \frac{2}{R_y + Z_B} (p \circ p_{MB}) \quad (3.21)$$

- Positive (top) y-direction particle velocity update equation:

$$u_y \circ u_{MT} = \frac{R_y - Z_T}{R_y + Z_T} (u_y \circ u_{MT}) + \frac{2}{R_y + Z_T} (p \circ p_{MT}) \quad (3.22)$$

- Pressure update equation (all elements):

$$\begin{aligned} p \circ p_M = p \circ p_M & - \frac{c^2 \rho \Delta t}{\Delta x} [u_x \circ u_{Mxp1} - u_x \circ u_{Mxp2}] \\ & - \frac{c^2 \rho \Delta t}{\Delta y} [u_y \circ u_{MyP1} - u_y \circ u_{MyP2}] \end{aligned} \quad (3.23)$$

where:	u_x	=	x-particle velocity element matrix
	u_y	=	y-particle velocity element matrix
	Z_L, Z_R, Z_T, Z_B	=	characteristic wall impedance values
	u_{xM}, u_{yM}	=	velocity masks, non-boundary velocity updates
	p_{Mx1}, p_{Mx2}	=	pressure masks, non-boundary velocity updates
	p_{My1}, p_{My2}	=	pressure masks, non-boundary velocity updates
	$u_{ML}, u_{MR}, u_{MT}, u_{MB}$	=	velocity masks, boundary velocity updates
	$p_{ML}, p_{MR}, p_{MT}, p_{MB}$	=	pressure masks, boundary velocity updates
	p_M	=	pressure mask, all pressure updates
	u_{Mxp1}, u_{Mxp2}	=	velocity masks, all pressure updates

Non-rectangular topologies can be efficiently modeled in two or three dimensions while allowing for any number of obstacles within the acoustical space. A comprehensive toolbox exploiting this ability is detailed in the following sections including the validation of the model's accuracy.

3.3 Simulation toolbox overview

Due to the lack of a publically/commercially available low-frequency acoustic simulation toolbox with adequate and appropriate flexibility, a bespoke toolbox was developed in MATLAB using the FDTD modeling techniques highlighted in the previous section. The initial aim of the toolbox was to provide a means of visualizing low-frequency behavior in small closed spaces due to any number of placed sources using real-time animations, overall sound-pressure distribution plots and frequency/time domain analysis. Upon achieving these goals, the toolbox was expanded to include a wealth of functions specifically tailored to the needs of this research. The key features of the toolbox are discussed in the forthcoming sections. An in-depth discussion on using the toolbox can be found in Appendix B.

3.3.1 Core simulation code

Equations 3.17 – 3.23 were converted into MATLAB code to serve as the core of the FDTD toolbox. As half time steps are not practical in digital computing, particle velocity and pressure updates were performed on whole time step intervals, with particle velocity updates always preceding pressure updates. This process is carried out within a *for loop*, operating using the variable *tnum*, which runs from 1 to the number of time steps needed to reach the desired simulation time (Fig. 3.9). Only the most recent matrix values for particle velocity and pressure values are saved, keeping memory requirements to a minimum.

Obstacles placed within the room do not necessarily have the same absorptive properties as the boundary surfaces within the space. Consequently, the boundary condition masks detailed in Section 3.2.6 must be split into two masks: one for room boundaries and one for obstacle surfaces. This is achieved by examining if a boundary in the mask is the first (or last) along a certain dimension. If so, it is considered a room boundary and is placed into the relevant mask. Otherwise, the element is considered an obstacle surface and is placed into the obstacle mask (see CD-ROM included with Appendix A for source code).

The distinction between room boundaries and obstacles allows for separate absorption coefficients for each variety of surface (and in each direction). This requires separate particle velocity boundary condition update equations (Fig. 3.9, lines 19-29).

Since all sound sources have predefined sound pressure signals, the corresponding pressure elements must be updated both with the predefined signal and with the resulting pressure response due to its surroundings. This is handled by updating the source pressure elements with the predefined signal (Fig. 3.9, lines 32-35) and then updating the source pressure elements along with all other pressure elements using the normal update equation (Fig. 3.9, lines 38-39).

```

1      % loop to update the velocities and pressures over the time steps, n
2 - for n = 1 : tnum
3     % update the non-boundary condition nodes for particle velocity
4     ux(uxM) = ux(uxM) - uCx*(p(pMx1) - p(pMx2));
5     uy(uyM) = uy(uyM) - uCy*(p(pMy1) - p(pMy2));
6
7     % update the velocity at the left wall
8     ux(uML) = ((Rx - ZL)/(Rx + ZL))*ux(uML) - (2/(Rx + ZL))*p(pML);
9
10    % update the velocity at the right wall
11    ux(uMR) = ((Rx - ZR)/(Rx + ZR))*ux(uMR) + (2/(Rx + ZR))*p(pMR);
12
13    % update the velocity at the bottom wall
14    uy(uMB) = ((Ry - ZB)/(Ry + ZB))*uy(uMB) - (2/(Ry + ZB))*p(pMB);
15
16    % update the velocity at the top wall
17    uy(uMT) = ((Ry - ZT)/(Ry + ZT))*uy(uMT) + (2/(Ry + ZT))*p(pMT);
18
19    % update the velocity at the left obstacle
20    ux(uMLI) = ((Rx - ZLI)/(Rx + ZLI))*ux(uMLI) - (2/(Rx + ZLI))*p(pMLI);
21
22    % update the velocity at the right obstacle
23    ux(uMRI) = ((Rx - ZRI)/(Rx + ZRI))*ux(uMRI) + (2/(Rx + ZRI))*p(pMRI);
24
25    % update the velocity at the bottom obstacle
26    uy(uMBI) = ((Ry - ZBI)/(Ry + ZBI))*uy(uMBI) - (2/(Ry + ZBI))*p(pMBI);
27
28    % update the velocity at the top obstacle
29    uy(uMTI) = ((Ry - ZTI)/(Ry + ZBI))*uy(uMBI) + (2/(Ry + ZBI))*p(pMBI);
30
31    % set the pressure at the source location(s)
32 - for z = 1 : nS
33     p(sgrid(z, 2), sgrid(z, 1)) = ...
34         source(z, n) + p(sgrid(z, 2), sgrid(z, 1));
35   end
36
37   % update the pressure at all nodes
38   p(pM) = p(pM) - pCx*(ux(uMxp1) - ux(uMxp2))...
39     - pCy*(uy(uMyp1) - uy(uMyp2));
40 end

```

Fig. 3.9 Core FDTD toolbox simulation code in MATLAB (two-dimensional case)

3.3.2 MATLAB graphical user interface (GUI) programming

A graphical user interface (GUI) for the FDTD simulation toolbox was designed using the GUIDE function in MATLAB. GUIDE allows a GUI to be created in a visual manner where text boxes, drop-down menus, plotting axes and pushbuttons are placed directly on a visual layout. The function converts the layout design to an appropriate source code structure, allowing for the necessary code to be inserted for each GUI object.

The FDTD simulation toolbox GUI allows all features and functions of the toolbox to be accessed within a single window, without any need to alter the source code (Fig. 3.10). The features include spatial topology configuration, source signal definition, data analysis and auralization.

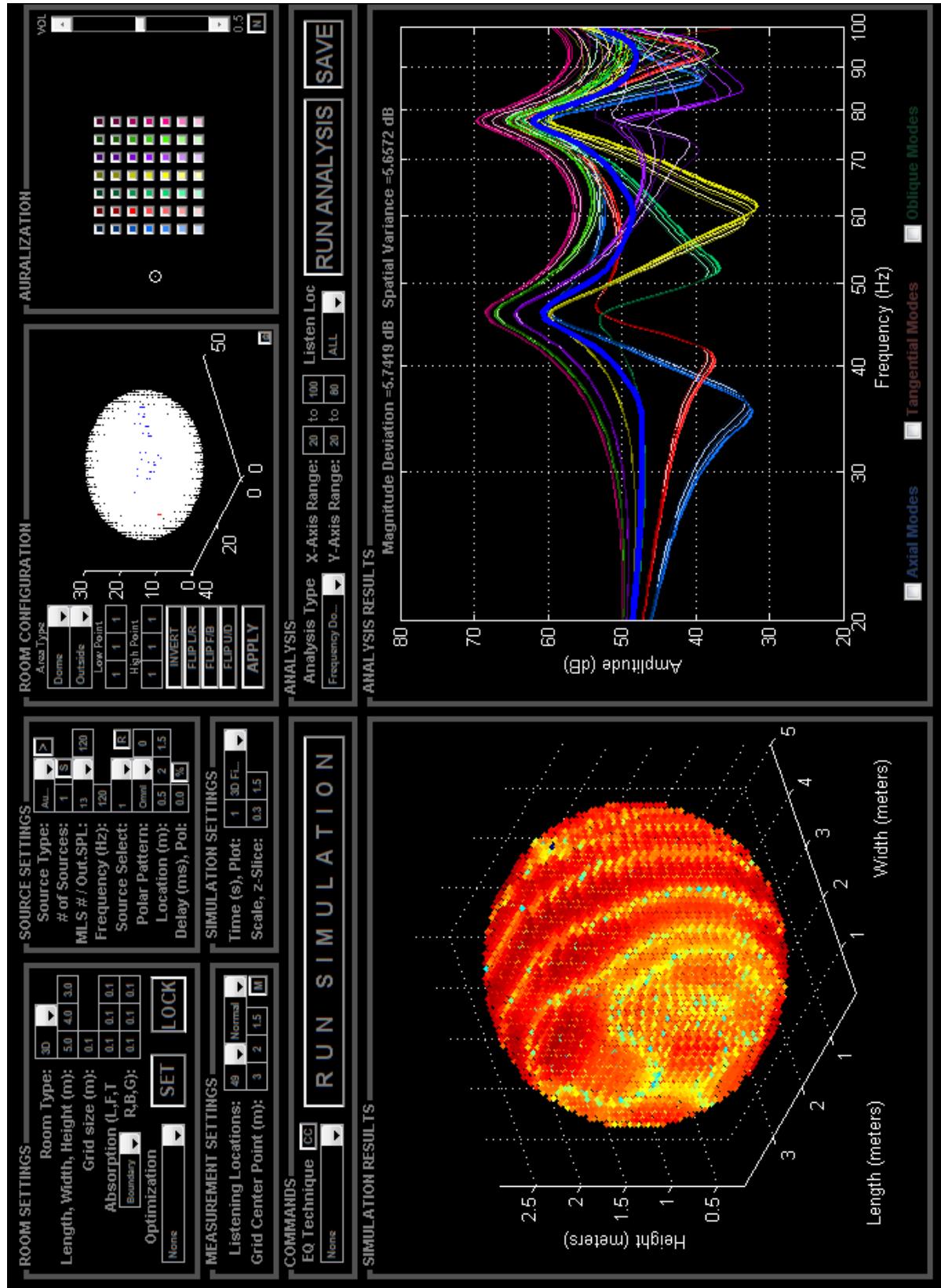


Fig. 3.10 FDTD simulation toolbox GUI in MATLAB

3.4 Simulation validation

It is of utmost importance to verify that the simulation toolbox exhibits a high degree of accuracy for low-frequency acoustical modeling. There are a number of methods for validation, including comparisons to theoretical room-mode predictions (both spectrally and spatially), real-world measurements and previously published results. These methods are considered in the following sub-sections, followed by a general discussion on the overall accuracy of the FDTD simulation toolbox.

3.4.1 Theoretical modes

The first step in validating the output of the simulation is to compare the simulated room-modes to those predicted theoretically. A virtual rectangular space of dimensions 5 m x 4 m x 3 m was configured with a single omnidirectional subwoofer in the corner and a twenty-five point listening grid. All surfaces were set to 10% absorption and an MLS signal was used to measure the response at the listening locations. A frequency response plot was generated with the theoretical modes included as colored vertical lines, where axial modes are blue, tangential modes red and oblique modes green (Fig. 3.11).

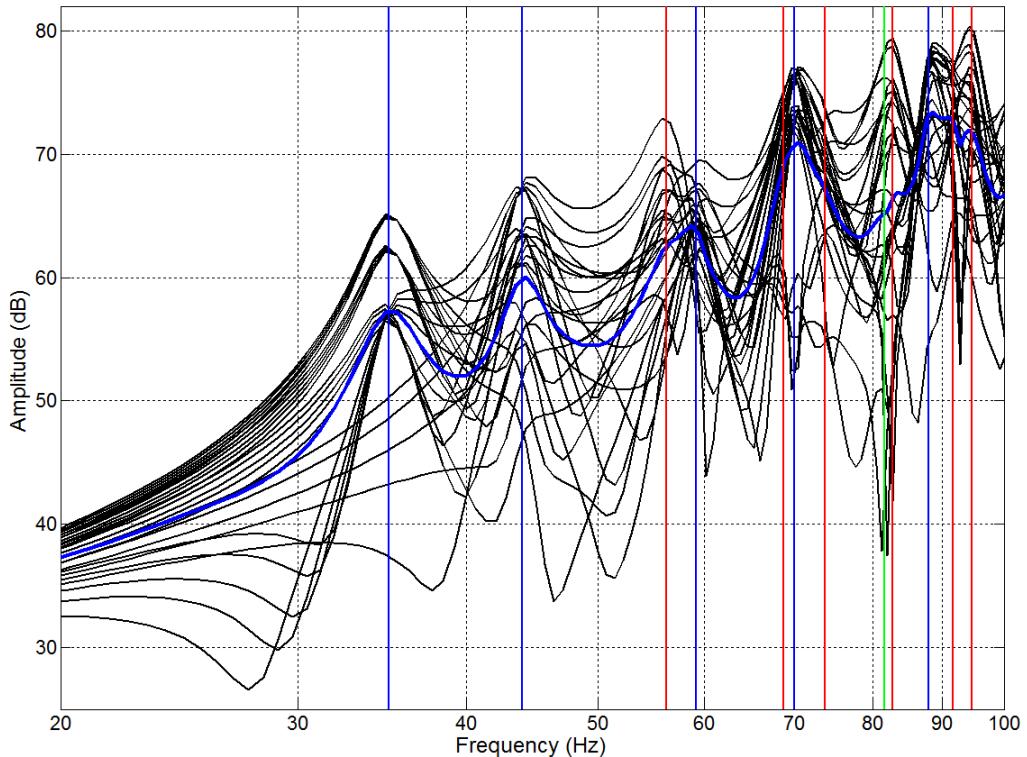


Fig. 3.11 Simulated frequency responses over a 25-point listening grid (solid blue line = room average response) in a 5 m x 4 m x 3 m room with theoretical modal frequencies displayed as vertical lines (axial = blue, tangential = red, oblique = green)

The theoretical and simulated modal comparison indicates that the FDTD toolbox is predicting modes at the correct spectral locations, where the axial modes tend to have more influence over the frequency response than the tangential and oblique modes.

Having verified the spectral accuracy of the simulated room-modes, it is necessary to examine the spatio-pressure distribution at various room-modes and compare them to the theoretical distributions, as predicted by the room-mode calculator function of the toolbox. Three modal

values were chosen for this comparison: 70.0 Hz (axial mode), 112.4 Hz (tangential mode) and 163.2 Hz (oblique mode). The theoretical spatio-pressure distributions comparisons to the simulated distributions are shown in Figs. 3.12 – 3.14.

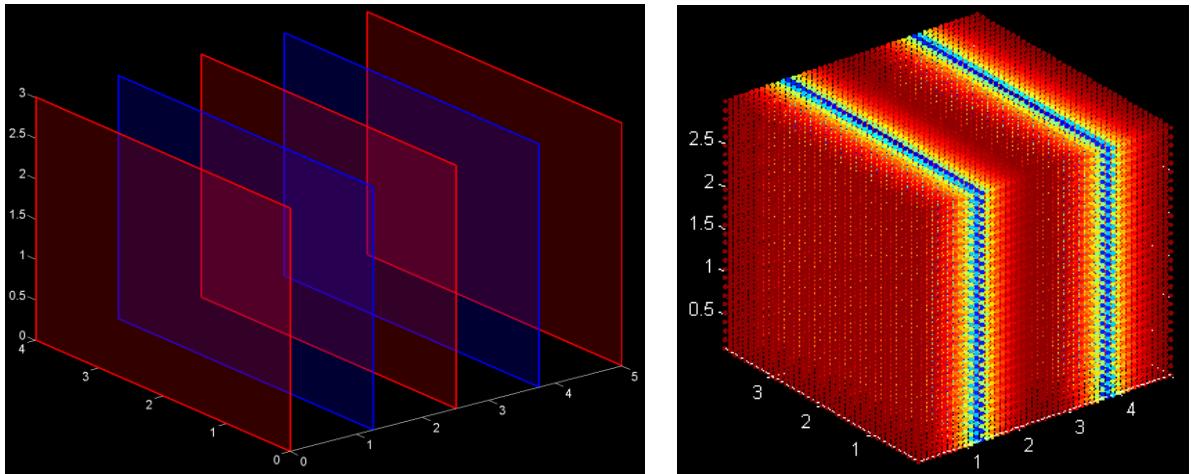


Fig. 3.12 Theoretical (left) and simulated (right) axial mode (2, 0, 0) spatio-distribution

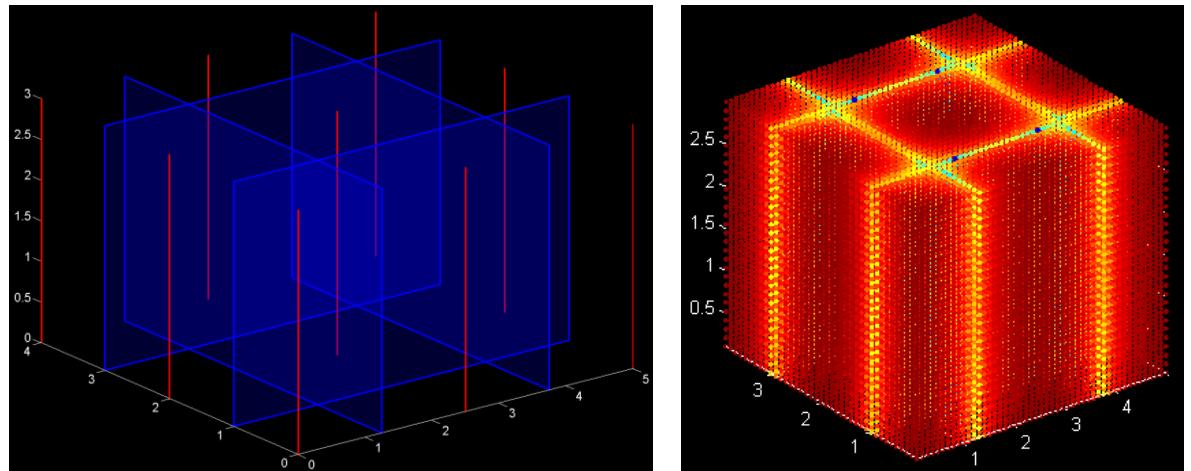


Fig. 3.13 Theoretical (left) and simulated (right) tangential mode (2, 2, 0) spatio-distribution

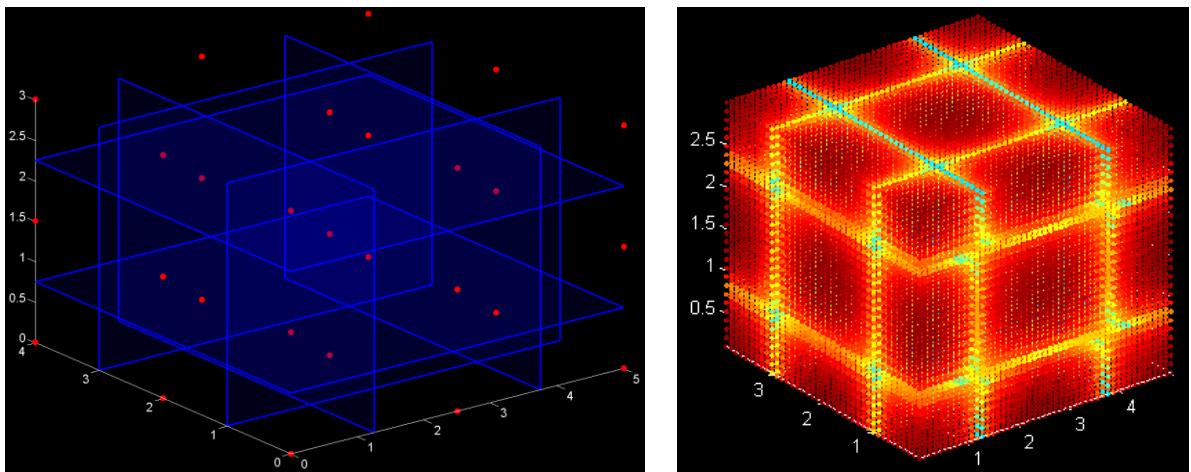


Fig. 3.14 Theoretical (left) and simulated (right) oblique mode (2, 2, 2) spatio-distribution

3.4.2 Lab measurements

While the simulation results agree nicely with theory, it is also necessary to compare simulations to real-world measurements. To accomplish this, measurements were taken in the Audio Research Laboratory (ARL) listening space. The measurements were taken using Clio 8.5 measurement software [47] with an Audiomatica MIC-01 measurement microphone. The ARL's rectangular dimensions are 6.05 m x 5.79 m x 2.80 m with a large cabinet structure of dimensions 1.73 m x 0.33 m x 1.56 m at the rear of the room. Nine listening locations were arranged in a square formation with 1.0 m spacing in both horizontal dimensions. The listening grid center point was located at (3.50 m, 3.00 m, 1.50 m).

First, the system was configured to include a single subwoofer in the corner of the room (0.40 m, 0.40 m, 0.30 m) and then measured and simulated using an MLS. The resulting frequency responses for all nine listening locations were plotted and compared side-by-side to examine the simulation toolbox's accuracy (Fig. 3.15).

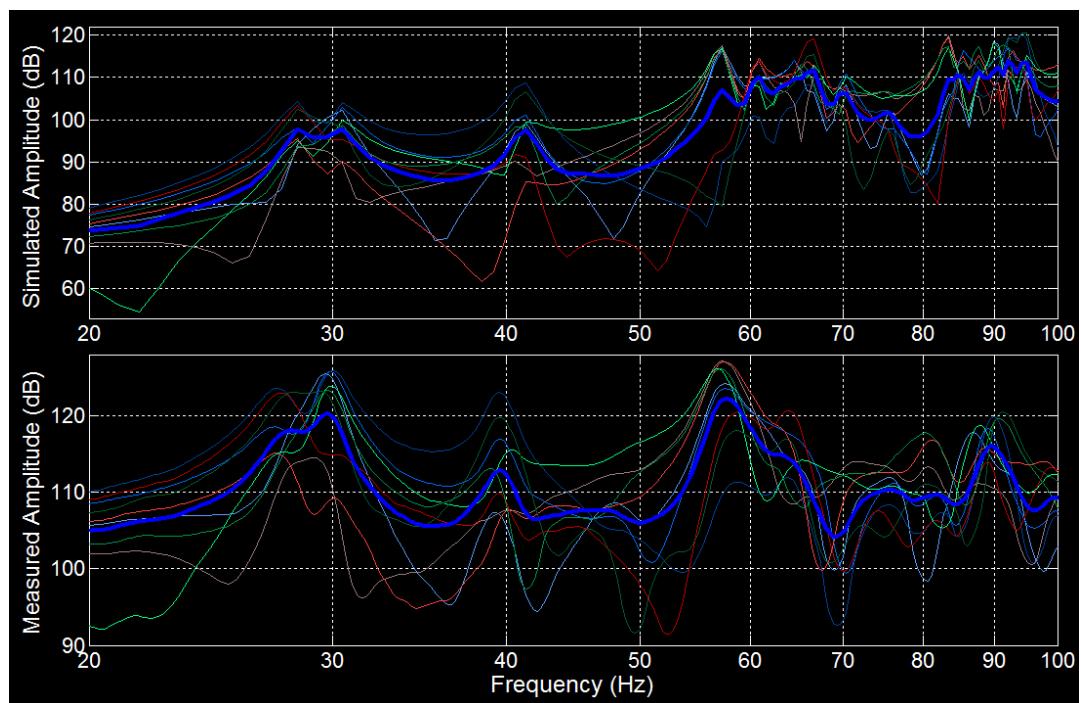


Fig. 3.15 Simulated (top) and measured (bottom) frequency responses for a single subwoofer system with room corner placement (solid blue line = room average)

The general trends of the frequency responses show similarities, especially at low frequencies. Inaccuracies can be attributed to a number of factors. First, the simulation does not include any reactive surfaces or frequency dependent absorption coefficients. Although the side walls and floor of the ARL are thick concrete, the ceiling consists of a series of porous acoustical tiles that tend to resonate when excited. Also, the simulation operates using an approximate bandpass response characteristic for the virtual subwoofers. In reality, the subwoofers used for the measurements do not exhibit a flat response over the central subwoofer range which likely colors the responses. Lastly, subwoofers are simulated as point sources, disregarding any diffraction effects of the enclosure. This simplification adds a slight inaccuracy to the toolbox, although if thorough simulations are required, an enclosure can be included.

In addition to the single room corner subwoofer system, a second subwoofer can be added in the opposite room front corner and then simulated, measured and compared (Fig. 3.16).

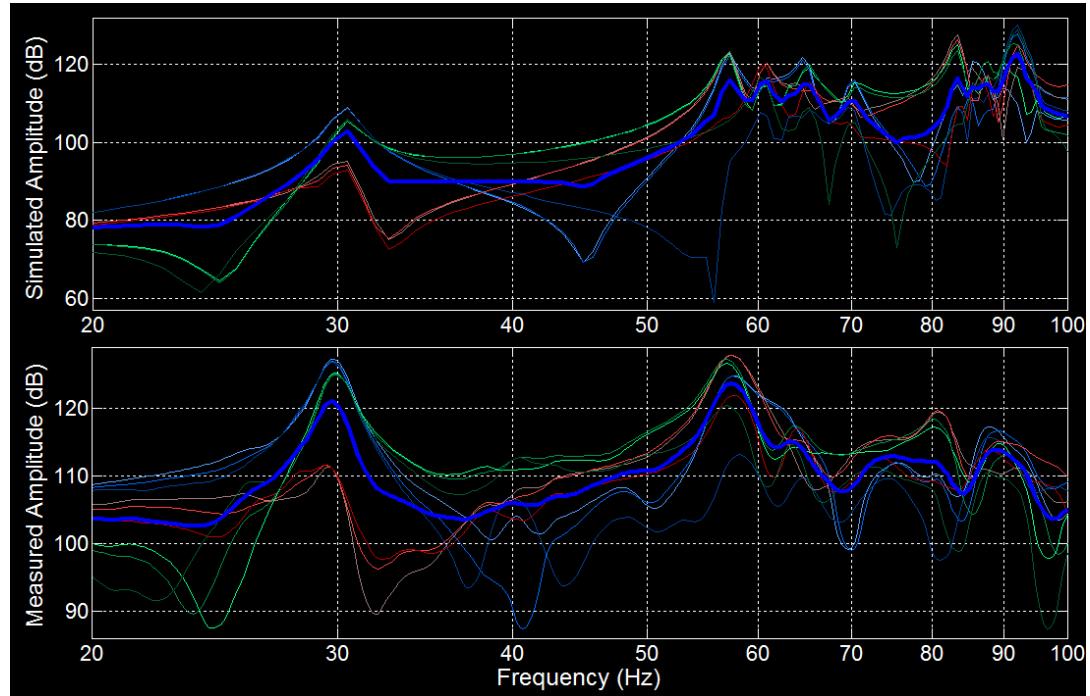


Fig. 3.16 Simulated (top) and measured (bottom) frequency responses for a dual subwoofer system with front room corner placement (solid blue line = room average)

The dual subwoofer system shows approximate agreement between simulations and measurements, where the general trends of the responses at each listening location are exhibited in each plot.

Despite the above mentioned inaccuracies, the comparison shows enough agreement between simulations and measurements to conclude that the toolbox is capable of delivering an approximate solution to various acoustical scenarios. As the toolbox was developed as a design aid instead of a high-accuracy simulation procedure this is acceptable since the toolbox predicts the strongest axial room-modes to within 1-2 Hz of the measured room-modes. These are the key frequencies that must be addressed using a correction procedure (to be developed using the toolbox), so it is important that the virtual and real environments are in agreement.

3.4.3 Previously published results

While comparisons to theory and real-world measurements provide clear indication if a simulation is accurate, it is also useful to compare simulation results to previously published results using different simulation methods. Welti performed a considerable amount of work concerning subwoofers in home theater applications where Welti simulated numerous room configurations for an Audio Engineering Society convention paper [17]. All published results compared to the FDTD simulation toolbox results are taken from this paper.

Welti employs a geometrical model of a rectangular topology in [17]. The relevant equations for this procedure are discussed in Section 2.2.2. Welti's configuration includes a room of dimensions 6.1 m x 7.3 m x 2.8 m with 5% wall absorption on all surfaces. A sixteen-point listener grid, centered in the room with 60 cm point-to-point spacing, was placed to monitor point-to-point response variation. The paper does not detail the height of said grid, so the center of the vertical dimension was utilized as the paper specifies the grid is perfectly centered in the room.

The first configuration explored is a subwoofer in the room corner. The frequency response was calculated from the measured responses from 0 – 100 Hz, in accordance with the data in [17]. The response comparison is shown in Fig. 3.17.

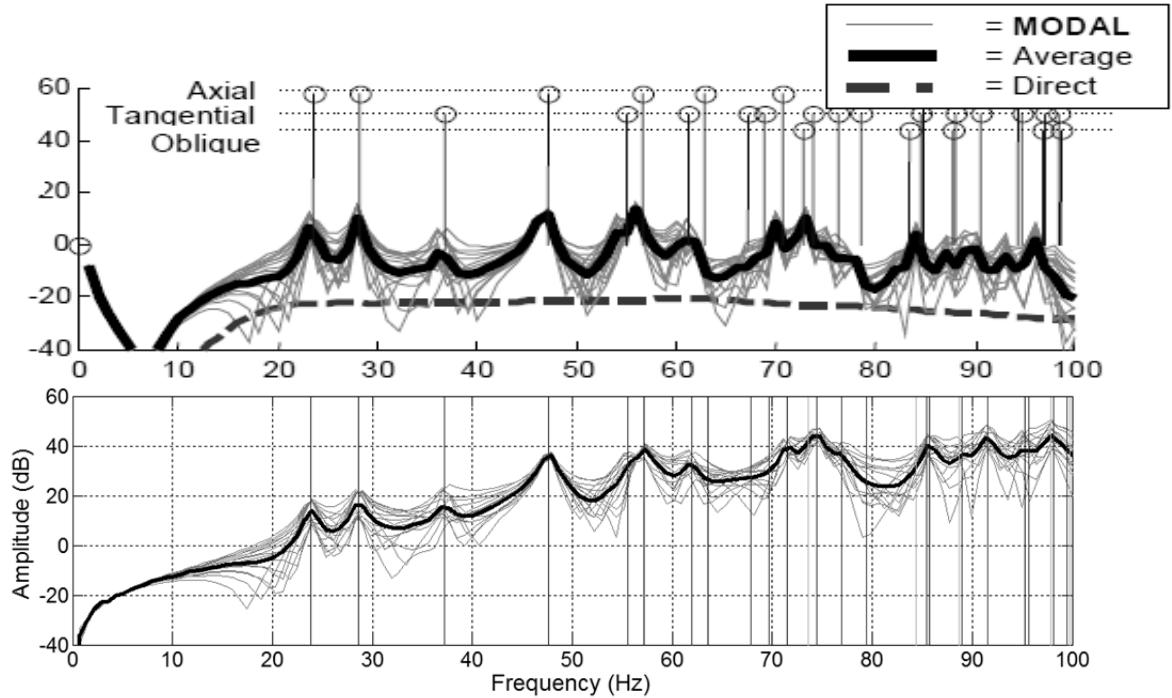


Fig. 3.17 Simulation results for a single subwoofer in a room corner from [17] (top) and the FDTD simulation toolbox (bottom)

The comparison between Welti's results and the FDTD toolbox results show close agreement across the entire spectrum, with the exception of the boost below 5 Hz in Welti's model, due to the mathematical model including the DC room-mode. The boost below 5 Hz occurs since below the lowest (non-DC) room-mode the sound field is determined by volume changes due to the source resulting in increasing room pressurization with decreasing frequency [5]. The FDTD toolbox does not display this behavior since the response is high-passed to approximate a subwoofer (although this function can be disabled, if necessary). This is also the cause of the gradually increasing sound pressure level with frequency. Next, results obtained when the subwoofer is shifted to a wall midpoint (on the ground) are compared (Fig. 3.18).

Again, the toolbox results agree closely with Welti's, with the exception of some vertical room-modes, likely due to the ambiguity of listening grid vertical height. The FDTD listener grid was placed perfectly in the center of the vertical dimension, the exact location of a number of room-mode nodal planes. Listening locations located along these nodes do not experience strong modal behavior, hence their absence in the FDTD simulation results.

Lastly, a multiple subwoofer system is investigated where a subwoofer is placed at wall midpoints along three of the four room walls (Fig. 3.19). The comparison of results from this multiple subwoofer configuration shows that the FDTD toolbox is in agreement with the results obtained by Welti's geometrical approach.

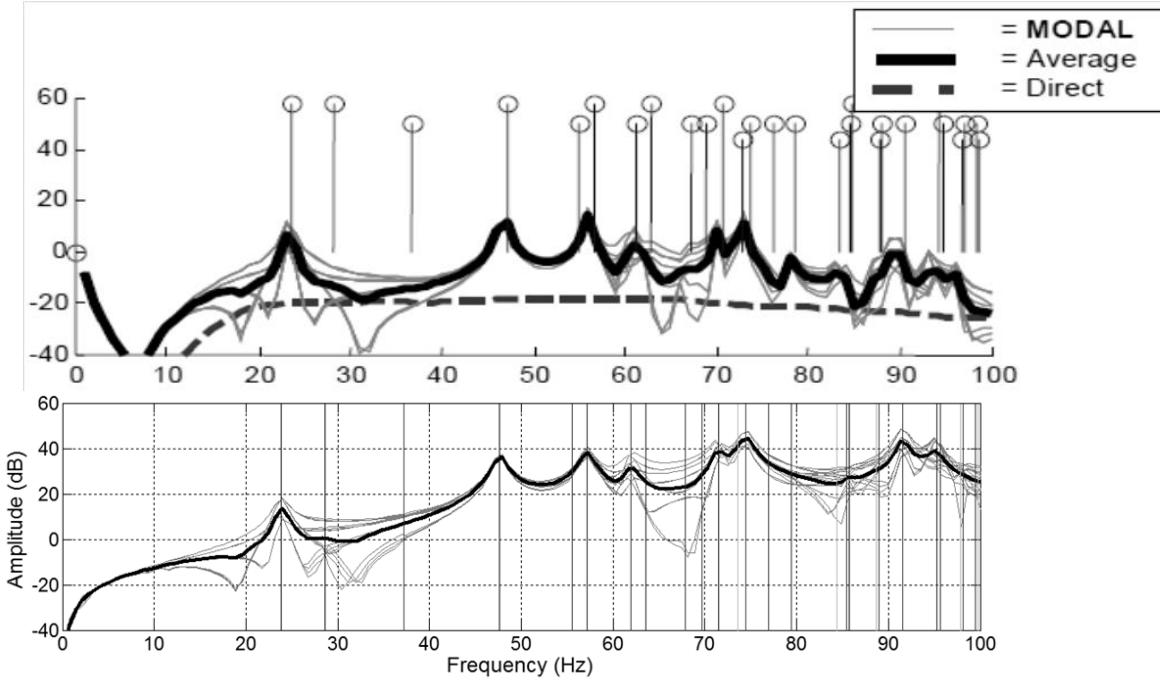


Fig. 3.18 Simulation results for a single subwoofer at a wall midpoint from [17] (top) and the FDTD simulation toolbox (bottom)

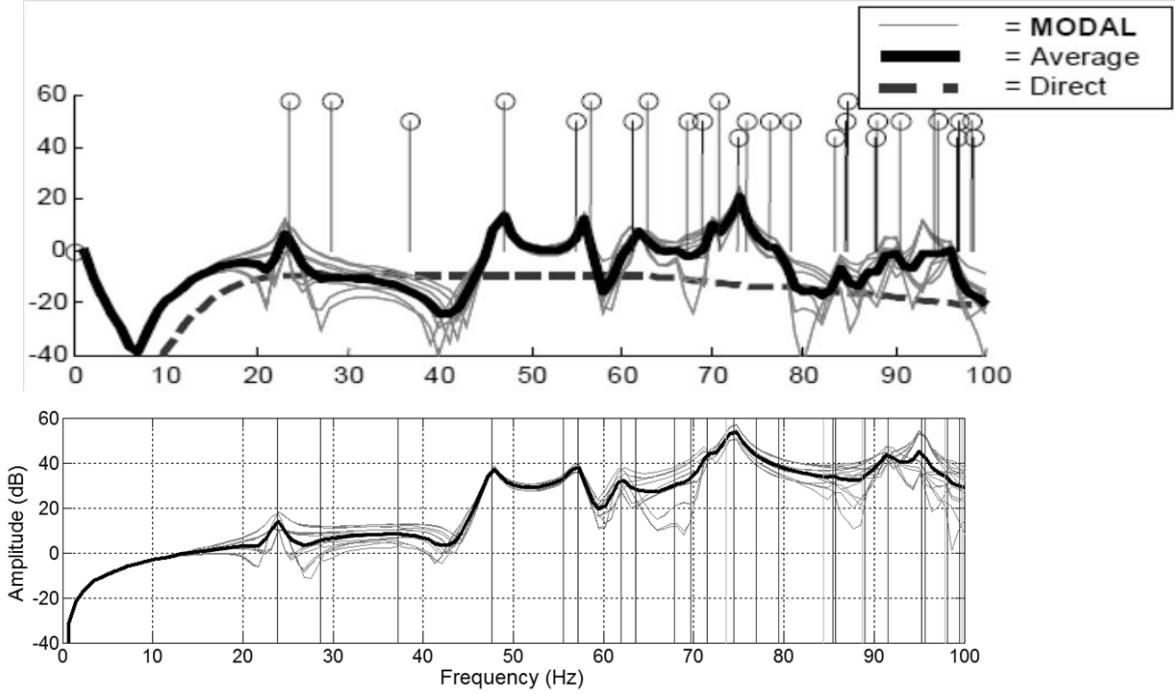


Fig. 3.19 Simulation results for three subwoofers at wall midpoints from [17] (top) and the FDTD simulation toolbox (bottom)

3.4.4 Discussion

Three different methods of simulation validation have been employed to examine the accuracy of the FDTD simulation toolbox developed for this research project. The comparison of simulated and theoretical spectral and spatial distribution of room-modes gives good agreement. In addition, the simulation results closely resemble those from published work, using different

acoustical modeling procedures. In comparing the simulation to real-world measurements, errors do arise. These errors can be attributed to approximations made within the simulation such as ignoring reactive surfaces, loudspeaker enclosures and high-complexly geometries. The measurements do show similar characteristics in frequency response to the simulations, notably at low-order axial modes which are the primary cause of significant spatial variation in small-rooms (and must be targeted using correction methods). This elevates the confidence that the simulation is sufficiently accurate for use primarily as a low-frequency room-mode correction methodology prototyping tool.

3.5 Chapter summary

There exist many adequate options for acoustical modeling. These options fit into three categories: geometrical, statistical and hybrid. Geometrical methods include image source and ray tracing, which exhibit good high-frequency accuracy while maintaining acceptable computational efficiency. Image source models lose efficiency as spatial topologies increase in complexity while ray tracing methods can have accuracy problems at low-frequencies since the wavelengths will be larger than the modeled space.

Statistical methods, on the other hand, provide very high low-frequency accuracy while also providing simple capabilities for tracking multiple sources and receivers. The three most common of these techniques are FEM, BEM and FDTD. While FEM and BEM operate in the frequency domain, which can result in transient accuracy issues, FDTD operates in the time domain yielding very accurate time and frequency domain simulations. The dominant disadvantage of these methods is the requirement of finer grid structures as frequency increases to avoid spatial aliasing. This produces a significant decrease in computational efficiency as the required upper simulated frequency limit increases.

Hybrid methods combine the best of both worlds, often utilizing a statistical method for the low-frequency band and a geometrical method for the high-frequency band. This ensures that the model produces acceptable low and high-frequency accuracy while keeping computational demands to a minimum.

The FDTD acoustical simulation method was chosen for this project since the research focuses only on low-frequency behavior yet requires high accuracy in both the time and frequency domains. The simulation method was linked into a GUI in MATLAB to allow for simple configuration of an acoustical space with extremely flexible data analysis functionality. The toolbox allows for animation/visualization of waveform propagation as the simulation is running as well as post-simulation data analysis with the aim to aid in the development of a prototype low-frequency room correction system.

The FDTD simulation toolbox results were validated using three methods. First, the simulated room-modes were compared to the theoretical predictions, both spectrally and spatially. Next, the simulation results were compared to real-world measurements using a number of subwoofer configurations. Lastly, the results were compared to published work which utilizes entirely different simulation methods.

The theoretical and published work comparisons show very strong agreement with the FDTD results both spatially and spectrally and with a number of different subwoofer configurations. The measurement comparisons also show similarities, although the results did differ from the FDTD data at certain points, especially at higher frequencies. This can be attributed to the

approximations within the simulation including details concerning the room topology, subwoofer enclosure and the acoustical properties of the wall materials.

The simulation toolbox is meant to provide approximate results for various acoustical scenarios to aid in the development of novel configuration and correction methods. Extremely high accuracy is not necessary. The validation process shows that the toolbox does agree with theory and different simulations used by notable researchers. Even though it does not agree perfectly with real-world measurements for the reasons stated, it can be taken as accurate for this work especially in the context of creating a well-defined virtual acoustic world for experimentation. A number of worked examples highlighting the broad usefulness of the toolbox are contained in Appendix B.

4 Conventional room-mode correction

The impact of room-modes on the spatiotemporal response across a wide listening area is discussed in detail in Chapter 2. Room-modes cause variable frequency responses between listeners which can prevent high-quality listening experiences for everyone. Methods aimed at room-mode suppression/reduction/correction have been researched extensively where many solutions achieve a considerable reduction in spatial variance.

This chapter highlights prominent methodologies for room-mode control, including passive and active approaches. In this work, passive approaches are classified as methods without any necessary electronic signal manipulation while active approaches are classified as those requiring some form of signal processing to achieve the desired results. The advantages and disadvantages of each approach are discussed in detail with many approaches being demonstrated using the FDTD simulation toolbox, which was described in Chapter 3.

4.1 Passive correction

The fundamental aspects of low-frequency response in small rooms involve the physical layout of the space. This includes the room dimensions, subwoofer location(s), listener location(s) and the quantity of subwoofers. Each of these aspects must be considered when designing a listening room containing an active sound playback system.

Since trial and error methods of room layout can take an unrealistically long time to carry out, metrics exist to highlight key attributes that can be objectively optimized to provide the best possible spatiotemporal response over the listening area. Many of these metrics (spatial variance, magnitude deviation, mean output level) are detailed in Chapter 2.

Published work on these topics is highlighted in the following sections. Room dimensions are explored including suggestions for optimal room-dimension ratios and how these ratios relate to the listening experience. Material properties are also discussed, concentrating on absorption levels and their relation to room-modes. Subwoofer to room coupling is explored, focusing on the effect of source/receiver placement and also on proposed single/multiple subwoofer configurations that can minimize spatial variance.

4.1.1 Room dimensions

A method of modal suppression often discussed in published work is room dimension optimization. Desirable room dimensions cause maximal spectral spacing between room-modes. This helps avoid modal groupings where adjacent modes have significant spectral overlap (or even share a center frequency) causing great irregularities over the low-frequency spectrum [48].

As room-modes are a function of any combination of room dimensions, problematic modal grouping is at its worst when room dimensions are integer multiples of one another. This results in each dimension's respective axial modes to perfectly correspond to those of the other dimensions. This extreme grouping causes large boosts in the frequency response at the modal frequencies, strongly coloring the room response.

With this in mind, a room quality index can be utilized, as presented by Walker in [48] (Eq. 4.1). This quality index takes into account the mean-square spectral distance between room-mode pairs.

$$QI = \frac{1}{N} \sum_{i=1}^N (f_i - f_{i+1})^2 \quad (4.1)$$

where: QI = mean-square room quality index
 N = number of modes within the frequency band
 f_i = modal frequency for the i^{th} mode

Using the mean-square room quality index, Walker proposed a room-dimension ratio criterion [48] (Eq. 4.2).

$$1.1^w/h \leq l/h \leq ^w/h - 4 \quad (4.2)$$

where: l, w, h = length, width, height of a room (m)

Walker suggests that this proposal permits smaller rooms with extremely even modal distributions. He claims that other room dimension ratio recommendations require unnaturally tall rooms (over four meters), even for small-sized listening rooms.

Other suggested criterions were generated through similar optimization calculations as with Walker's, but using total surface area and edge length along with room volume [49]. This is accomplished by determining the expected number of room-modes below a specified frequency (Eq. 4.3) [49]. The equation is commonly simplified to the form contained in Eq. 2.11. This modal density calculation requires an accurate calculation of the expected number of room-modes, therefore Eq. 4.3 is employed. The modal density (modes/Hz) is calculated by taking the derivative of Eq. 4.3 with respect to frequency and then carrying out the necessary computations (Eq. 4.4) [15, 49].

$$N(f) = \frac{4\pi V}{3c^3} f^3 + \frac{\pi S}{4c^2} f^2 + \frac{L_R}{8c} f \quad (4.3)$$

$$\frac{\partial N}{\partial f} = \frac{4\pi V}{c^3} f^2 + \frac{\pi S}{2c^2} f + \frac{L_R}{8c} \quad (4.4)$$

where: L_R = total length of all room edges (m)
 $\frac{\partial N}{\partial f}$ = modal density (modes/Hz)

Using this calculation technique, Milner arrives at the ratio of 1 : 1.186 : 1.439 for rectangular rooms, although he emphasizes that better modal spacing can be achieved in non-rectangular spaces such as rooms with at least one set of non-parallel walls [49].

One ratio appears more than any other with noticeably less experimental evidence to prove its superiority. This popular ratio is commonly referred to as the "Golden Ratio," which is based on the Fibonacci Sequence where each number (except the first in the sequence) is the sum of the previous two sequence numbers (1, 1, 2, 3, 5, 8, 13, 21, 34, ...). After progressing through the sequence (to around 377) a constant ratio appears between adjacent values. This ratio, ϕ (Phi), equals 1.61803 (or 0.61803 if the ratio is inverted) and is often suggested to exist in nature, art, music and architecture. Specifically, it has been observed that the cochlea in the human ear is in the shape of a "golden spiral," which is generated using the ratio, ϕ [50]. Whether this is a factor in preference to music composed using the golden ratio, has yet to be sufficiently proven.

Even with lacking scientific evidence for its value, the golden ratio appears in room acoustics discussions quite regularly. The golden ratio for a three-dimensional room is 1 : 1.62 : 2.62 [50]. Again, while this ratio is similar to the two objectively defined ratios presented in this section, little work has been done to prove the superiority of the golden ratio. In addition to the golden ratio, there are also “Silver” and “Bronze” ratios that are derived based on solutions to similar quadratic equations. (Eqs. 4.5 – 4.7) [51]. These ratios have yet to be proven as relevant to audio and acoustics. In recent years, claims of the golden ratio superiority have been disproven in select publications [51, 52].

$$x^2 - x - 1 = 0 \quad (4.5)$$

$$x^2 - 2x - 1 = 0 \quad (4.6)$$

$$x^2 - 3x - 1 = 0 \quad (4.7)$$

where: $\text{Eq. 4.5 roots} = \text{Golden Ratio (1 : 1.618)}$

$\text{Eq. 4.6 roots} = \text{Silver Ratio (1 : 2.414)}$

$\text{Eq. 4.7 roots} = \text{Bronze Ratio (1 : 3.303)}$

A function built into the FDTD simulation toolbox assists in room dimension optimization. This function takes in a range of possible room dimensions set by the user and performs frequency response measurements across a grid of listening locations for each dimension combination. Spatial variance is calculated for each permutation to determine the configuration resulting in the minimal spectral variance. Magnitude deviation and mean-square room quality index are also calculated and the corresponding best and worst case scenarios for these measures are determined. Often (but not always) the best-case is the same for all three values. It is important to note that spatial variance is sensitive to listening area placement. Moving the measurement point grid generally results in significant changes to the characteristics of a room dimension analysis plot.

A rectangular room of height 3.0 m was set up in the toolbox with a grid spacing of 10 cm. A 25-point listening grid was positioned in the center of the room with a height of 1.8 m with a single omnidirectional subwoofer placed at a room corner. The width and depth of the room were swept from 3.0 to 6.0 m, in steps of 10 cm. Upon completion of the simulation, a plot was generated showing spatial variance due to room width and depth (Fig. 4.1).

The spatial variance displayed in Fig. 4.1 was calculated from 20 – 120 Hz (subwoofer operating band). As modal frequencies are a function of room dimensions, the decrease in spatial variance with decreasing room dimensions can be clearly seen. In smaller rooms the lowest possible room-mode is at a much higher frequency than in larger rooms. Since the frequency band used to calculate spatial variance is fixed at 20 – 120 Hz and smaller rooms naturally have a larger portion of that analysis range sitting below the lowest mode (where spatial variance is minimal), small spaces are expected to exhibit lower spatial variance than larger rooms using this form of room dimension analysis (as seen in Fig. 4.1).

The dimension optimization indicates that spatial variance is minimized for room dimensions 4.1 m x 3.9 m x 3.0 m. This corresponds to a dimension ratio of 1.00 : 1.30 : 1.37, and is similar to the ratios devised by Milner [49]. Magnitude deviation calculations give approximately the same ratio, while the room quality index indicates the best configuration is 5.6 m x 5.1 m x 3.0 m, giving a ratio of 1.00 : 1.70 : 1.87. This ratio agrees perfectly with the relationship proposed by Walker [48].

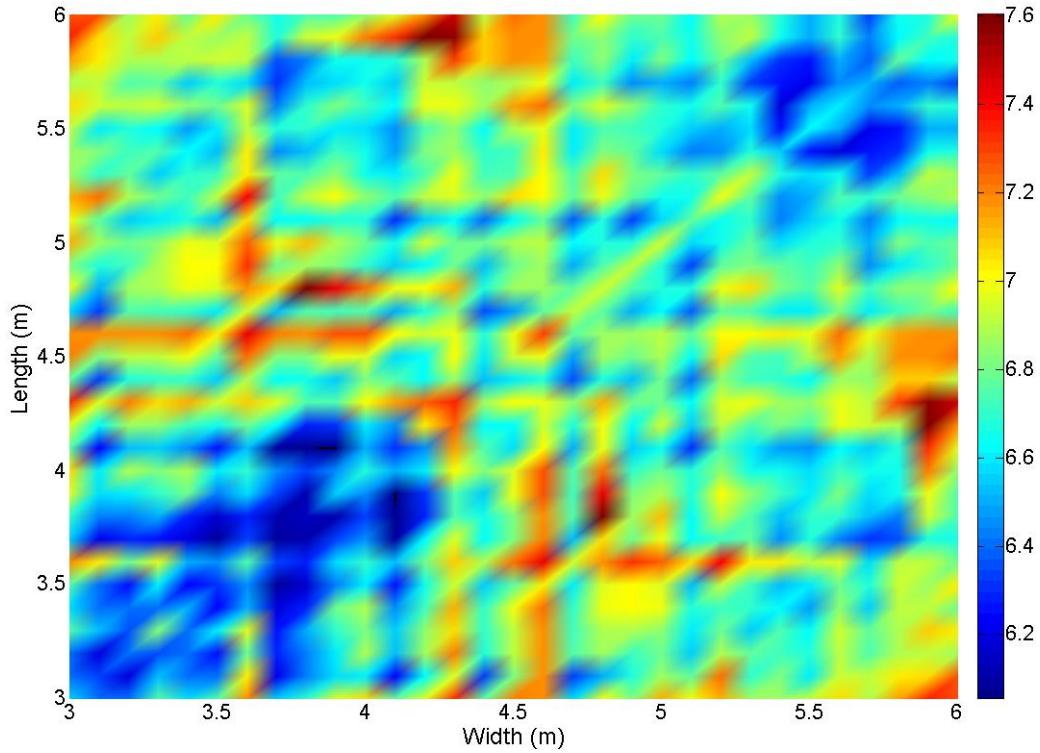


Fig. 4.1 Spatial variance (dB) versus room dimensions (room height fixed at 3 m)

Although spatial variance is largely dependent on source-to-room coupling, a few simple rules can be deduced from the theoretical and simulated results. First, room dimensions should not be integer multiples of one another. Doing so results in significant modal overlap and strong resonances. Even though spatial variance decreases with room size, this should be avoided if system efficiency is important. If the bulk of the subwoofer operating band (20 – 120 Hz) falls below the lowest room-mode then this range will not benefit from acoustical reinforcement since the wavelengths will be significantly larger than the room dimensions. In addition, smaller rooms may not be practical for systems designed to be listened to by many individuals at once, simply due to spatial restrictions.

It is important to note that although spatial variance can be reduced with appropriate room dimensions, the reduction is not significant (around 1 – 2 dB). Supplementary techniques must be utilized if further spatial variance reduction is required; many of such options are discussed later in this section. Despite the lack of spatial variance reduction, a carefully designed room avoids overly strong low-frequency resonances, which ensures no strong coloration of the room response due to the room geometry.

4.1.2 Absorption

Increasing surface absorption in a room is one technique to mask modal problems. As most wall materials exhibit similar absorption at low-frequencies (anywhere from 1 – 10%), surface absorption can be increased by adding soft, porous materials such as foam or fiberglass to the walls of a room. These materials effectively dissipate acoustical energy as the air particles encounter a significant amount of resistance when moving through the material causing friction, thus converting the acoustical energy to heat [2].

The decreased reflections from the walls limit the buildup of standing waves and cause the modes to exhibit a lower Q, resulting in less noticeable resonances due to increased modal

overlap (Fig. 4.2). While Fig. 4.2 highlights a noticeable decrease in sharp room-modes as absorption is increased tenfold, spatial variance only decreases by a marginal amount (5.0%). Even though the acoustic space exhibits significantly fewer sharp resonances with added wall absorption, a strong variance still exists between listeners.

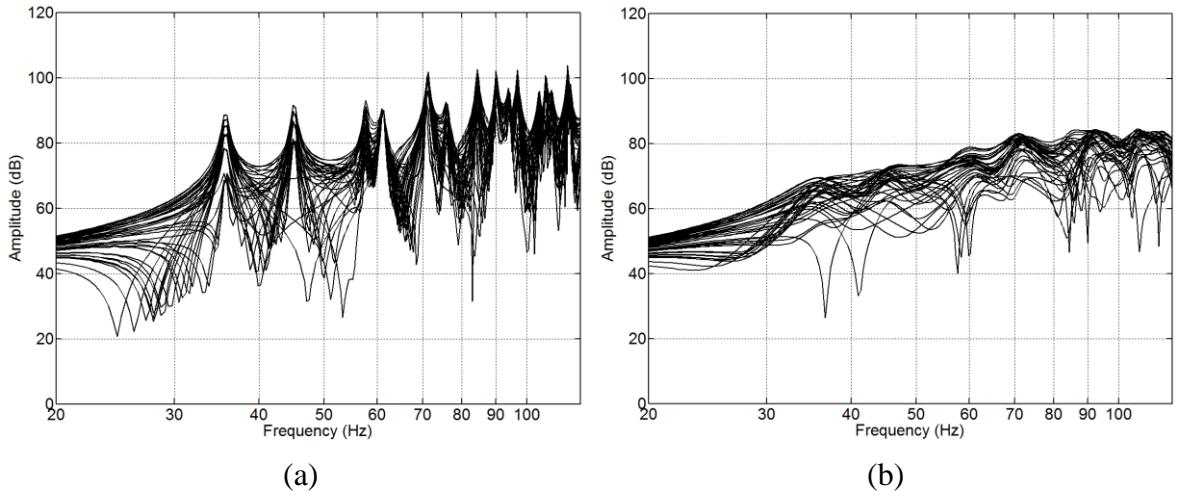


Fig. 4.2 Frequency response simulations over a 25-point listening grid for surface absorption of (a) 2% and (b) 20%

Foam-based absorbers are regularly used to reduce high-frequency reflections (most often in recording studios and high-grade listening rooms), but can be difficult to implement for lower frequencies since their size is dictated by the wavelengths in the target frequency band. As the absorbers operate by attenuating particle velocity, logic dictates that they must be placed at particle velocity maxima. At the room boundaries, particle velocity is theoretically zero; therefore this form of absorption would be very weak at this location. In order to maximize efficiency, absorbers must be placed a quarter-wavelength from the surface, at the particle velocity maximum for the axial room-modes [2].

While this is no problem for high-frequencies, long wavelengths at low-frequencies require these absorbers to intrude by over two meters into a room! This is impractical, but is nonetheless implemented for certain applications (primarily anechoic chambers). Disregarding the unrealistic size requirements of the passive absorbers, passive absorbers do not contribute greatly to spatial variance reduction, thus cannot be considered a relevant solution.

Bass-traps are commonly implemented to provide additional absorption at key room-modes (not over a wide spectral range). A variety of bass-traps are available in practice, including porous absorbers and resonating absorbers. These approaches each have their advantages and disadvantages, as discussed in [2, 53, 54]. Overall bass-traps can suppress the most problematic modal resonances, but cannot address spatial variance over the entire subwoofer operating band unless an unreasonable number of traps are included in a space.

4.1.3 Single subwoofer placement

The low-frequency room response can be significantly altered with subwoofer placement. Most commercially available subwoofers have omnidirectional polar patterns where the subwoofer operates as a pressure source. An omnidirectional subwoofer causes differing room responses depending on whether it is placed near a pressure node (pressure = minimum, particle velocity = maximum) or a pressure anti-node (pressure = maximum, particle velocity = minimum) [19]. When placed at a node, the subwoofer exhibits weak coupling with the room, causing the room-

mode to be minimally excited (Fig. 4.3a). The opposite is true when placement is near an anti-node (Fig. 4.3b) [2].

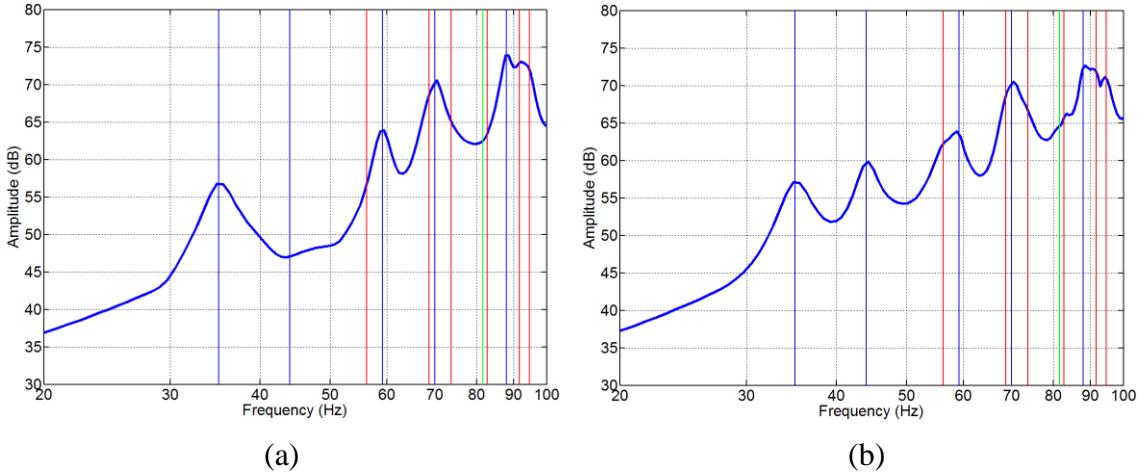


Fig. 4.3 Average frequency response over a 25-point listening grid with a single omnidirectional subwoofer at (a) (0.4 m, 2.0 m, 0.4 m) and (b) (0.4 m, 0.4 m, 0.4 m)

The suppressed mode shown in Fig. 4.3a corresponds to the (0, 1, 0) axial mode at 42.9 Hz. Using the room-mode spatio-distribution function from the FDTD toolbox (highlighted in Section 3.4.1) to assess the two source locations tested, it is clear that the corner location (0.4 m, 0.4 m, 0.4 m) is close to an anti-nodal plane for 42.9 Hz, while the central wall location (0.4 m, 2.0 m, 0.4 m) is directly on a nodal plane (Fig. 4.4). As expected, the corner location strongly excites the room-modes while the central wall location causes minimal excitation.

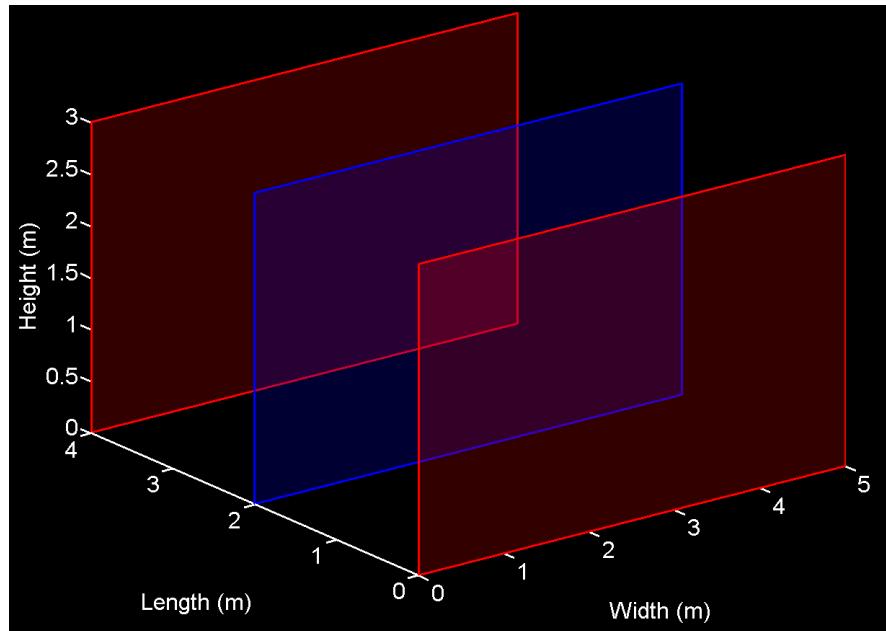


Fig. 4.4 Theoretical modal spatial distribution for the (0, 1, 0) mode in a 5 m x 4 m x 3 m rectangular room (red = antinodal plane, blue = nodal plane)

Considering the above results, placing the source directly in the center of the room should suppress many low-order axial modes that have antinodal points/planes at the room center. The low-order modes of (1, 0, 0), (0, 1, 0) and (0, 0, 1) all have been minimally excited due to the source's central placement, as displayed in Fig. 4.5.

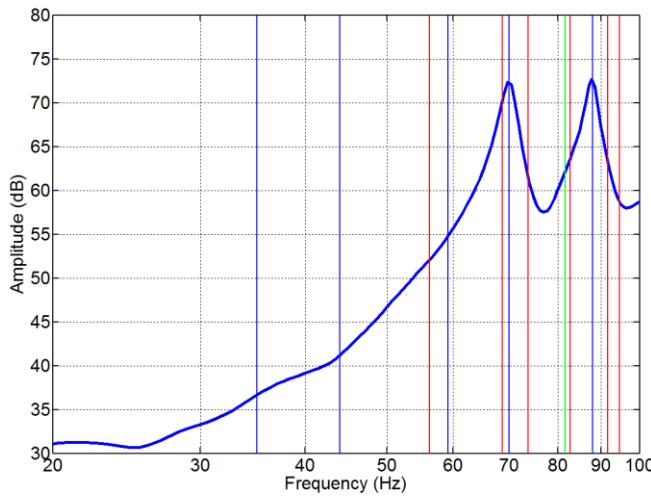


Fig. 4.5 Average response over a 25-point listening grid with a single omnidirectional subwoofer at the room center (2.5 m, 2.0 m, 1.5 m)

While central placement shows increased modal suppression, it is not a practical location for a subwoofer. Consequently, a single source position optimization routine is provided in the simulation toolbox, with spatial variance minimization in mind. For example, the position optimization routine can be carried out for placement along the y-axis wall with a distance range to the wall from 0.2 to 2.6 m (Fig. 4.6).

The single source placement optimization routine indicates, as discussed earlier, that central room placement results in the lowest spatial variance (although this is not a practical placement). Wall midpoint placement, on the other hand, is a reasonable source location with reduced spatial variance. It must be noted that nodal-source placement minimizes excitation of unwanted room-modes, thus the subwoofer output is not strongly reinforced by the room due to the low room-to-source coupling. As a result, the system is less efficient requiring higher subwoofer output capabilities; a fact that must be considered during the system design and configuration process.

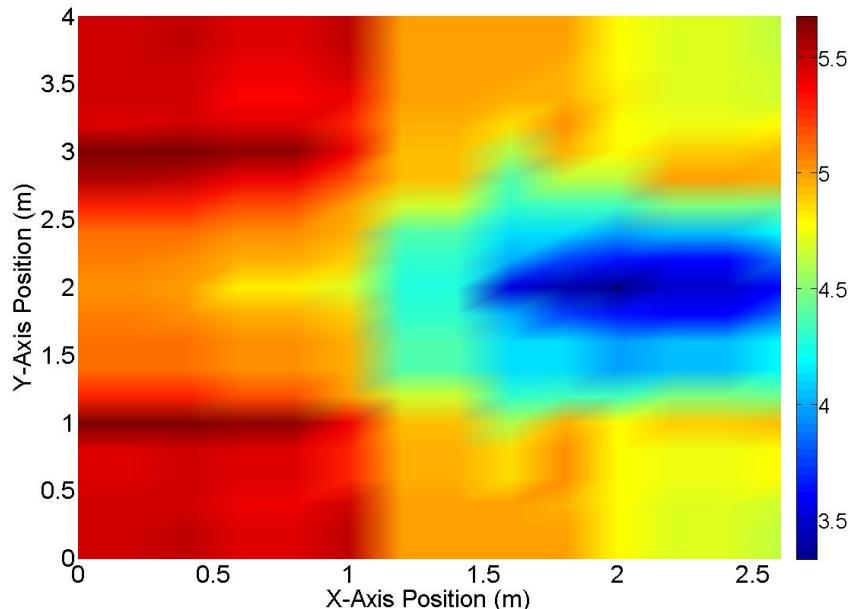


Fig. 4.6 Spatial variance (dB) for various subwoofer positions within a 5 m x 4 m x 3 m rectangular room (subwoofer height = 0.4 m)

4.1.4 Multiple subwoofers

More often than not, subwoofer placement is restricted by practicality (e.g. objects present in the room and aesthetics), where placement of a single subwoofer at the center of a rectangular room to optimally reduce resonances is unrealistic. A solution to this compromise scenario is to employ multiple subwoofers located at anti-nodal positions of opposite polarity. That is, one subwoofer is placed at a location corresponding to a positive pressure amplitude maximum while another is placed at a negative pressure amplitude maximum. The opposite polarity of anti-nodal positions causes destructive interference, minimizing excitation [2]. This placement gives similar results (Fig. 4.7a) to placing a single subwoofer at that mode's node (Fig. 4.3a), which utilizes a single subwoofer at a node to achieve modal suppression. Two more subwoofers can be placed at the other corners to help suppress the first-order mode along the width of the room (Fig. 4.7b) and thus achieve further spatial variance reduction.

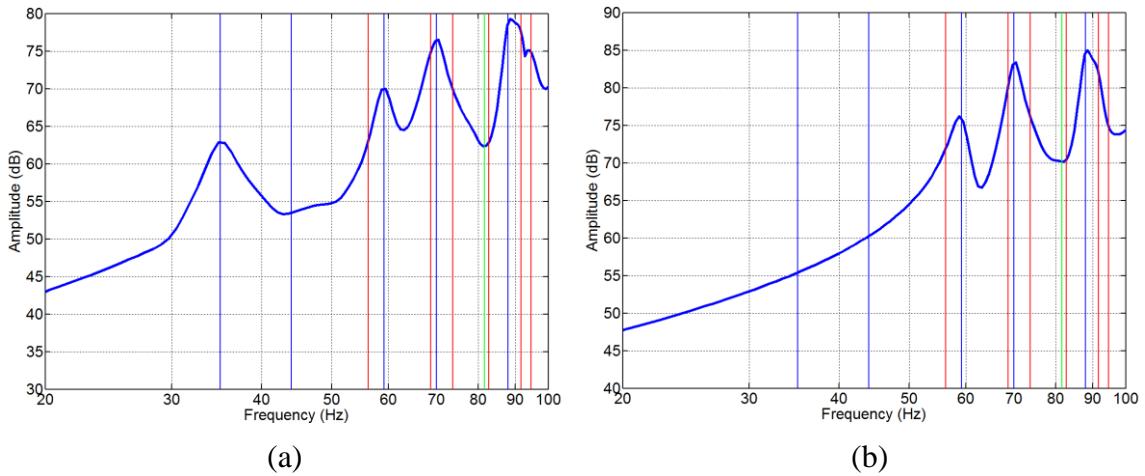


Fig. 4.7 Mean frequency response over a 25-point listening grid with (a) two omni sources at (0.4 m, 0.4 m, 0.4 m) and (0.4 m, 3.6 m, 0.4 m) and (b) four omni sources at (0.4 m, 0.4 m, 0.4 m), (0.4 m, 3.6 m, 0.4 m), (4.6 m, 0.4 m, 0.4 m) and (4.6 m, 3.6 m, 0.4 m)

Again, system efficiency must be considered, noting that sound pressure in the room is not linearly related to the number of sources. With each additional source, there exists more destructive interference which suppresses the problematic room-modes; hence, a doubling of sources in a room will not correspond to a doubling of sound pressure.

An efficiency/spatial variance compromise suggested in previous work [2, 19] employs four omnidirectional subwoofers placed at wall midpoints on the ground (Fig. 4.8). This placement follows the optimal placements indicated in Fig. 4.6, where modal suppression is maximized due to strategic nodal placement of the single sources and opposing anti-nodal placement for the two pairs of sources.

This configuration gives a spatial variance reduction of 74.3% (20 – 100 Hz), although the resulting frequency response is far from flat, exhibiting a strong resonance around 60 Hz, due to the first-order axial mode in the vertical dimension. This problem could in theory be eliminated using parametric equalization since each listening location would benefit equally from this global application of correction (although this falls under the “active” correction methods discussed in the following section). A possible downside to this approach, though, could be a smearing of the transient response, depending on the type and order of equalization utilized. Above 70 Hz the benefits of this form of passive correction are not clear, due to difficulties in suppressing tangential and oblique modes by source placement.

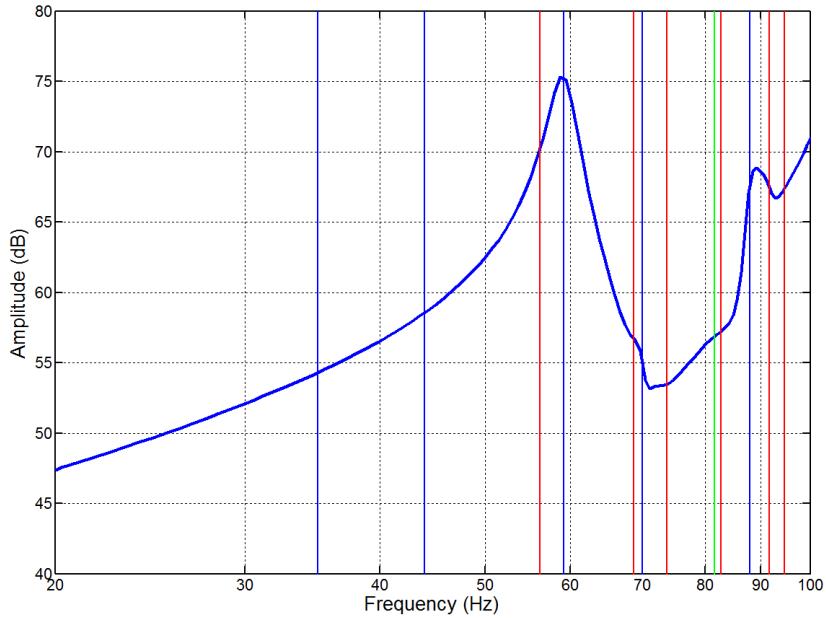


Fig. 4.8 Average frequency response over a 25-point listening grid with 4 omni sources at (0.4 m, 2.0 m, 0.4 m), (4.6 m, 2.0 m, 0.4 m), (2.5 m, 0.4 m, 0.4 m) and (2.5 m, 3.6 m, 0.4 m)

4.1.5 Passive correction conclusions

A number of passive techniques geared towards spatial variance minimization exist, ranging from physical room construction to single/multiple source placement. Room dimensions are crucial in determining the spectral modal distribution. Dimensions that are equal or integer multiples of one another share modal frequencies, causing strong room resonances which can prohibit a subjectively-pleasing acoustical environment. A number of dimension ratios have been previously suggested to give an even spectral distribution of room-modes to avoid this problem. This technique, however, does little to limit spatial variance and therefore cannot be used on its own to achieve an accurate spatiotemporal response across a large listening area.

Similarly, surface absorption is used to shape a room's acoustical characteristics but does not fully suppress all room-modes, rather widening resonances, reducing wall reflections and suppressing just a few dominant modes. Absorption at low-frequencies is impractical as the absorbers must intersect points where particle velocity is maximum; usually a quarter wavelength from a surface for axial modes. This requires unrealistically long absorbers, thus is not a practical solution in most cases.

Single/multiple source positioning is utilized for spatial variance reduction by placing sources at either mode nodal points to avoid the excitation of a particular room-mode or conversely placing sources at opposing mode anti-nodal points to achieve the same effect as with nodal placement, but with destructive interference. A four-subwoofer system with wall midpoint placement is recommended in the literature [2, 19] as a suitable solution to the problem, which has been verified with the FDTD toolbox to reduce spatial variance by over 70% as compared to a single corner subwoofer placement. While this is a strong step towards delivering an equal listening experience over a wide listening area, the multiple source configurations are only efficient in suppressing axial modes and therefore are only beneficial in the lower range of the subwoofer band. In the case of the 5 m x 4 m x 3 m room, this corresponds to modal suppression below 70 Hz. Spatial variance is largely unchanged above this point due to the un suppressed tangential and oblique modes.

4.2 Active correction

As discussed in the previous section, many forms of low-frequency room-mode control involve the physical configuration of a room and sound reinforcement system. While adjusting the system layout can improve spatiotemporal accuracy over a wide area, it can only work up to a point. Beyond that, active correction is necessary.

Active correction can involve a number of different techniques. One method that is regularly used involves generating an inverse filter based on a measured impulse response at a primary listening location to flatten the frequency response. Other methods expand on this principle to include multiple measurement points to generate the inverse filter, thus addressing the response over a wider area.

Multiple-point techniques have been expanded upon in certain published work to allow for adaptive filtering where the filter coefficients are continuously updated to compensate for environmental changes. These systems keep a room response as close to optimal as possible regardless of any changes to the room and/or system. Most of these techniques are predominantly focused on the frequency magnitude response to avoid filter instability, but other techniques exist that address phase as well. Phase correction can be performed independently of frequency magnitude correction and vice versa.

Aside from filter-based correction, a number of alternative approaches have been heavily researched including subwoofer polar pattern control, active absorption and radiation resistance-based control. The most common forms of active correction are discussed in this section, highlighting the advantages and disadvantages of each while providing simulation examples, where appropriate.

4.2.1 Static single-point equalization

A linear time-invariant (LTI) system is characterized by both an impulse response (time domain) and a complex frequency response (frequency domain) which are related by a Fourier transform pair (Eqs. 4.8 & 4.9) [55].

$$H(\omega) = \int_{-\infty}^{\infty} h(t)e^{-j\omega t} dt \quad (4.8)$$

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega)e^{j\omega t} d\omega \quad (4.9)$$

where:

$H(\omega)$	=	complex frequency response
$h(t)$	=	impulse response

The complex frequency response can be converted to polar form to give a clear definition of the magnitude and phase response (Eq. 4.10).

$$H(\omega) = |H(\omega)|e^{j\phi(\omega)} \quad (4.10)$$

where:

$ H(\omega) $	=	magnitude response
$\phi(\omega)$	=	phase response

Traditionally, a graphic or parametric equalizer is inserted into the reproduction chain in order to attenuate large peaks in the magnitude response to give a roughly flat (linear) response. This technique, while simple to implement, requires manual fine tuning which prohibits precise correction results. Additionally it has been claimed, although not supported by empirical data, that a flat response is not perceptually satisfactory in many cases as individuals are used to hearing the natural acoustical characteristics of a space [70].

To circumvent manual system correction, an inverse filter based on a transfer function measurement can be generated. The transfer function, commonly represented in the z-domain, is expressed by its zeros and poles (Eq. 4.11). The transfer function is said to be minimum phase when all zeros lie within the unit circle and, in such cases, the complex response of the system can be directly related using the Hilbert transform. If this is the case, precise system correction (in terms of magnitude and phase) can be performed with a single inverse filter since it will be LTI (Eq. 4.12) [56].

$$H(z) = \frac{B(z)}{C(z)} \quad (4.11)$$

$$A(z) = \frac{C(z)}{B(z)} \quad (4.12)$$

where:	$H(z)$	=	transfer function
	$B(z)$	=	transfer function zeros
	$C(z)$	=	transfer function poles
	$A(z)$	=	inverse filter for response correction

If a stable filter cannot be generated due to a non-minimum phase transfer function it is still possible to create an inverse filter, but it can only be based on the minimum phase components of the transfer function. The excess phase components can then be dealt with separately [56].

Inverse filter generation by Eqs. 4.11 and 4.12 is the simplest form of single-point equalization, commonly used in commercially available home theater systems. This method, in theory, works well for correcting acoustical anomalies at the primary listening location to give a flat (or any other desired) response.

Room-to-listener coupling is not identical for each location in a room. Therefore, single-point equalization cannot benefit a wide listening area. While the primary location may be corrected properly, other locations can deviate greatly from a flat (or the desired) response. This gives little (or no) reduction in spatial variance (Fig. 4.9).

Although the primary listening location (red line in Fig. 4.9) is equalized to give a flat response all other locations still exhibit wildly differing responses. Since the target point naturally exhibits a low-amplitude response around 55 Hz, this has to be compensated within the inverse filter, thus all other points in the room show a strong unnatural resonance at 55 Hz after equalization. Clearly, this system does not contribute to spatial variance minimization.

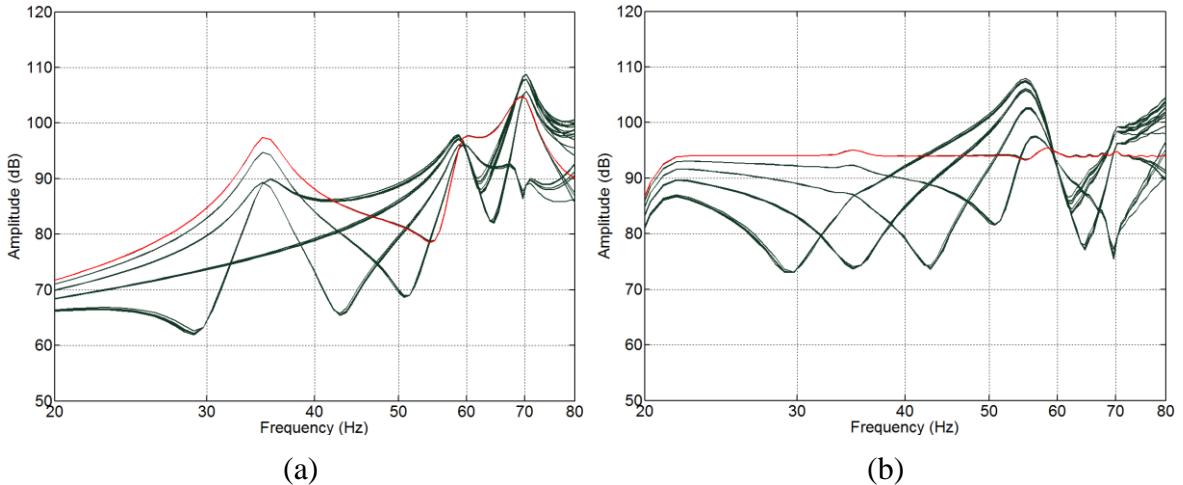


Fig. 4.9 Simulated frequency responses over a 25-point grid with a single subwoofer at (0.4 m, 2.0 m, 0.4 m) with (a) no equalization and (b) single-point equalization at the red line

4.2.2 Adaptive single-point equalization

Static single-point equalization techniques operate as linear, time-invariant systems. In other words, correction is applied to a signal in a predefined manner. Adaptive filters, on the other hand, are linear, time-varying systems, adjusted with a time-variant error signal. The error signal is usually defined as the difference between a predefined target response and the measured response at a single listening location [56].

A correction filter is constructed from this signal at each time window in order to minimize the error. This minimization routine is generally performed using a least-mean squares (LMS) algorithm, although this is not the only technique available [57]. It has been argued, although without much objective evidence, that the LMS criterion is not optimal in regards to psychoacoustics [58]. Regardless of the algorithm utilized, the generalized process of determining the error signal is shown in Fig. 4.10 [59].

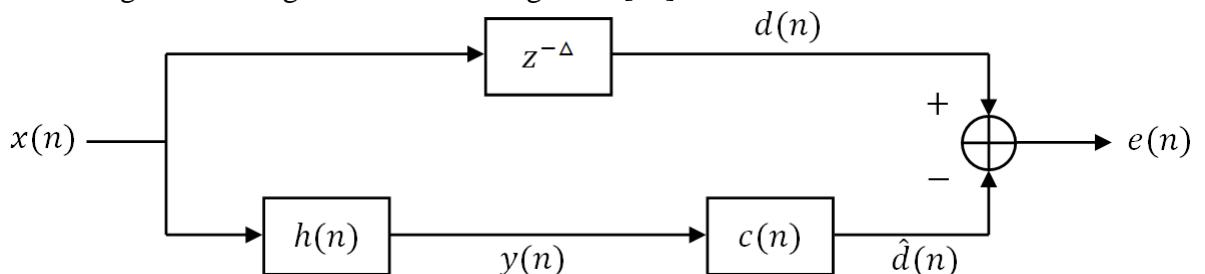


Fig. 4.10 Adaptive single-point equalization error signal calculation structure [59]

where:

$x(n)$	=	source signal
$h(n)$	=	impulse response of the correction filter
$c(n)$	=	impulse response of the system
$z^{-\Delta}$	=	modeling delay
$e(n)$	=	error signal
$y(n)$	=	equalized source signal
$d(n)$	=	target signal (delayed source)
$\hat{d}(n)$	=	measured signal

Adaptive filters are often of the finite-impulse response (FIR) variety. FIRs are used due to the simplicity of defining their response, where the filter's resolution is governed by the number of coefficients [56]. The fewer coefficients, the shorter the time window, resulting in less time delay and greater time domain accuracy, but with a reduced frequency resolution. The filter frequency resolution is given by:

$$f_{res} = 0.89f_s/N_c \quad (4.13)$$

where:	f_{res}	=	filter frequency resolution (Hz)
	f_s	=	sample rate (Hz)
	N_c	=	number of filter coefficients

As with static single-point equalization, single-point adaptive equalization provides accurate correction at the target location, but can cause other points to exhibit far worse spatiotemporal responses than before correction. It has been found that single-point equalization corrects over an area corresponding to a circle with a 2.8 cm radius (for correction up to 200 Hz) [59]. The area of correction is determined by one-sixth the shortest wavelength in the correction band. As source-to-listener coupling is not identical at all locations, correction is only relevant extremely close to the target location [56].

4.2.3 Multiple-point equalization

One solution to the problems encountered with single-point equalization is to add more measurement points within a listening area. The equalization process is roughly the same as with single-point equalization, but with the addition of extra measurements. These additional measurements are used within an expanded single-point equalization structure to determine their individual error signals. The goal with the adaptive filter variety is to use a procedure to minimize all error signals. The ideally flat response probably will not be achievable since the system is attempting to find a solution for all error signals; however the listening locations are likely to exhibit much less spatial variance. The process for finding the error signals in adaptive multiple-point equalization procedures is shown in Fig. 6.11 [59].

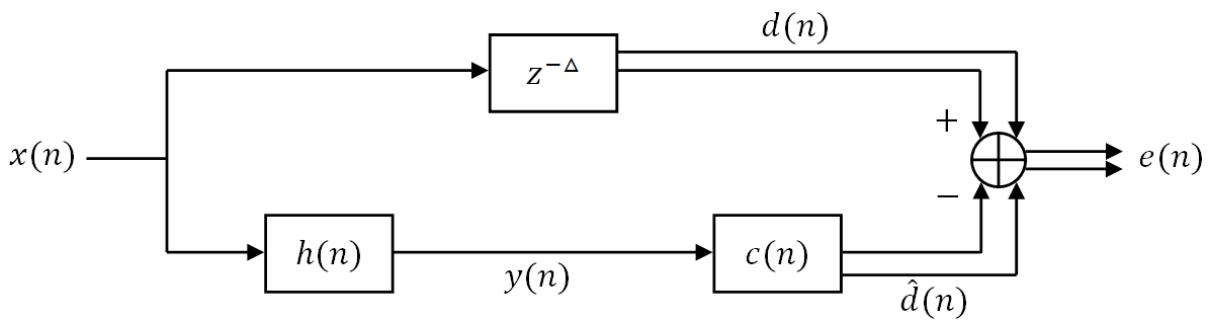


Fig. 6.11 Adaptive multiple-point equalization error signal calculation structure [59]

This form of filter generation raises the issue of how to weight individual error signals within the minimization process. The simplest solution equally weights each response. This assumes each listening location is distinctly different from the others and must be handled individually [59]. In recent years, however, a considerable amount of research has focused on multiple-point equalization using specialized weighting functions [60 – 64].

The bulk of these explorations analyzes measured responses and determines similarities between each individual response (either in the time or frequency domain). The procedure is performed over multiple frequency bands, allowing for the formation of similarity clusters in each band. The error signal weighting is therefore based on the number of responses included in each cluster. More responses in a given cluster cause that cluster to have a higher weighting than clusters with fewer members. The advantage to this process is that the most common frequency responses within a listening area are concentrated on while responses that occur only at single-points do not have a huge effect on equalization. The discussed research [60 – 62] mentions that it is possible to have low correlation between all measurements, preventing the formation of clusters. In this case, equalization utilizes traditional multiple-point static/adaptive equalization where all points are equally weighted.

This weighting procedure is discussed in earlier work [65], but without clustering. In this case individual responses are weighted based on an LMS criterion. It is suggested that this system can be used even if only a single equalization filter is available to the system.

It is important to note that a simple averaging of measured room responses used to generate a single inverse filter is unlikely to produce a reduction in spatial variance. This is due to low correlation between listening location responses, therefore an averaging of responses will not produce a filter that specifically addresses any one location, let alone them all. What is likely to happen is that the responses will remain equally divergent from the target response, although the room average will be very close to the target (Fig. 4.12).

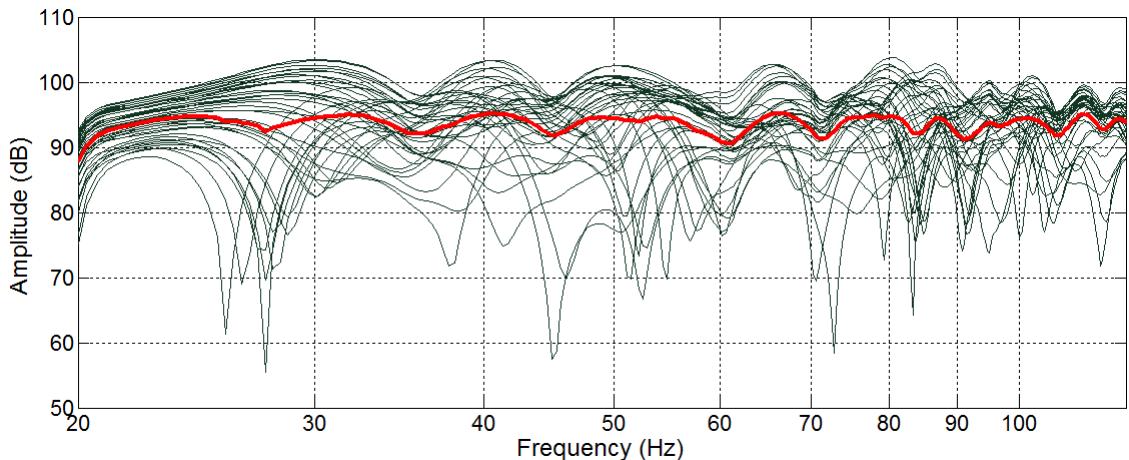


Fig. 4.12 Simulated equalization results using a static multiple measurement point, single filter room average approach (red line = equalized room average)

4.2.4 Polar pattern control

All low-frequency room-mode correction methods discussed to this point assume omnidirectional subwoofers. The non-directional radiation of these sources excites room-modes equally when placed near a corner. High modal excitation can be seen as advantageous for a number of reasons. First, subwoofers can deliver high-amplitude low-frequency content due to natural reinforcement from the room acoustics, as alluded to in Chapter 2. This phenomenon, known as the Waterhouse effect, becomes stronger as a source approaches more room boundaries.

In addition to increased output at modal frequencies, omnidirectional sources perform well below the lowest room-mode due to their direction-independent radiation which serves to pressurize a room when no modal reinforcement is present [5]. However, as demonstrated in

previous sections, a problem with omnidirectional sources is the substantial spatial variance below the Schroeder frequency [13]. There exist a number of approaches regarding spatial variance minimization where this section focuses on a more radical approach using low-frequency directivity control.

Subwoofer directivity is achieved using the principles presented by Olson concerning gradient loudspeakers [66]. Gradient loudspeakers utilize the techniques of microphone polar pattern control, but applied in reverse. These methods (with the exception of the zero-order variety) call for two or more spaced drive-units within each source to achieve the desired directionality. A number of configurations are presented by Olson which provides a useful tool set for low-frequency polar pattern control.

The zero-order gradient source is the building block for all higher-order gradient sources. This consists of a single drive-unit which radiates energy equally in all directions (Fig. 4.13).

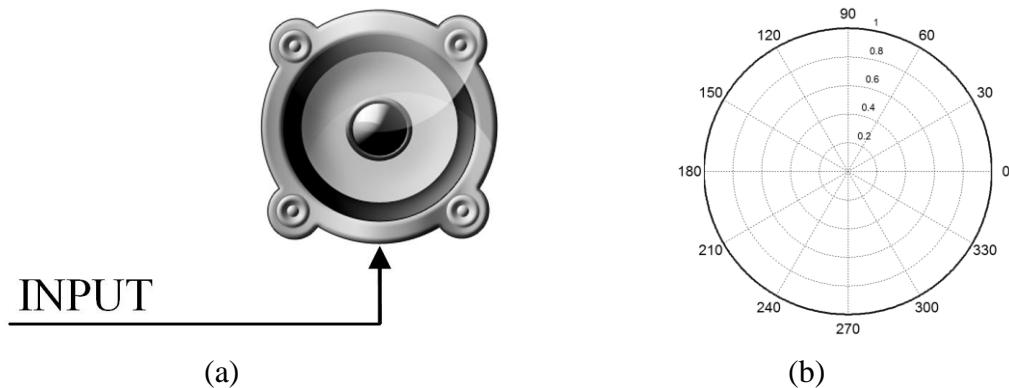


Fig. 4.13 Zero-order gradient source (a) configuration and (b) polar pattern

First-order gradient sources combine two zero-order sources, one of reverse polarity (Fig. 4.14). This configuration's polar pattern is highly-dependent on the physical separation of the two sources (Eq. 4.14).

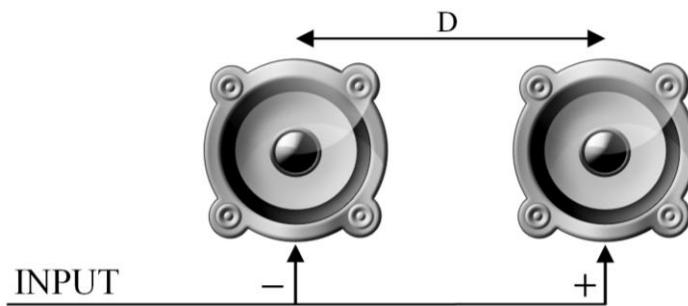


Fig. 4.14 First-order gradient source configuration (dipole)

$$R_\theta = \sin\left(\frac{kD}{2} \cos\theta\right) \quad (4.14)$$

where:	R_θ	=	system response at θ (rad) from on-axis
	k	=	wave number
	D	=	source spacing (m)

Spacing at a quarter-wavelength of the target frequency results in a dipole pattern, but spacing at a full wavelength gives a four-lobed pattern (Fig. 4.15).

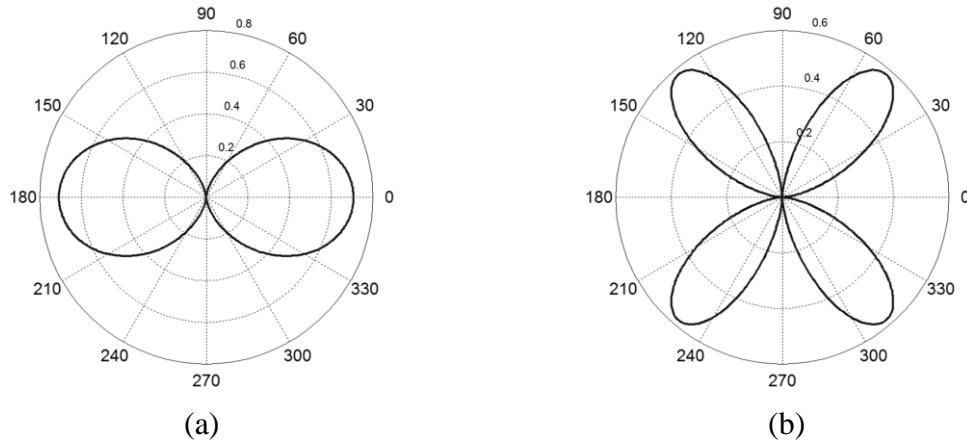


Fig. 4.15 First-order gradient source (dipole) polar pattern with drive-unit spacing at
(a) $\frac{1}{4}$ wavelength and (b) full wavelength

The dipole first-order gradient source can be adjusted to give a cardioid pattern. This involves adding electronic delay to the second drive-unit which directly corresponds to driver spacing (Fig. 4.16). Again, the polar pattern is highly-sensitive to driver spacing (Eq. 4.15) as quarter wavelength separation gives a cardioid pattern, but full wavelength separation gives a dipole pattern rotated 90° off-axis (Fig. 4.17).

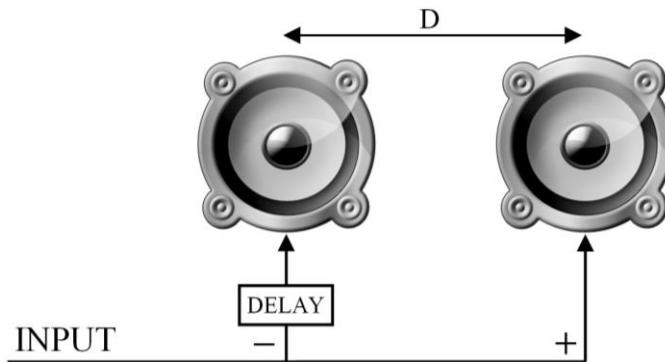


Fig. 4.16 First-order gradient source configuration (cardioid)

$$R_\theta = \sin\left(\frac{kD}{4} + \frac{kD}{4} \cos\theta\right) \quad (4.15)$$

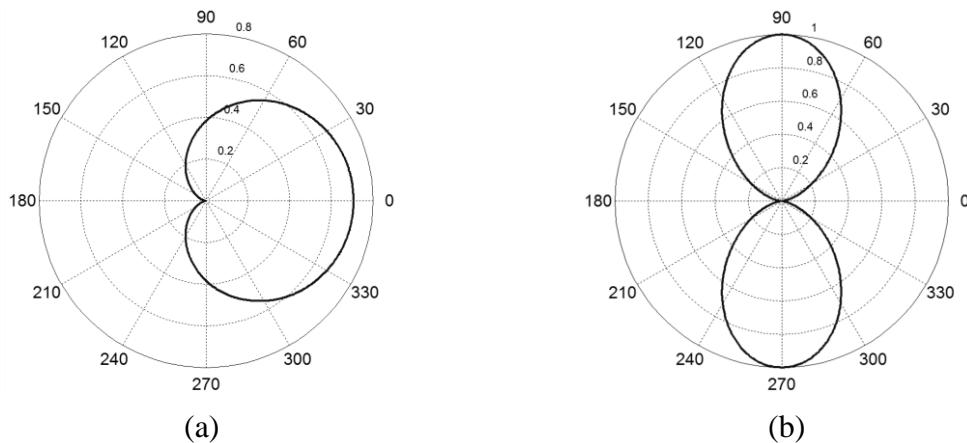


Fig. 4.17 First-order gradient source (cardioid) polar pattern with drive-unit spacing at
(a) $\frac{1}{4}$ wavelength and (b) full wavelength

Second-order gradient sources are formed using two dipole first-order sources and placing them together with a physical separation and an electronic delay on the second first-order source directly corresponding to the separation distance (Fig. 4.18). The polar pattern is described by Eq. 4.16 [66] and, like with other gradient configurations, is largely dependent on source spacing (Fig. 4.19).

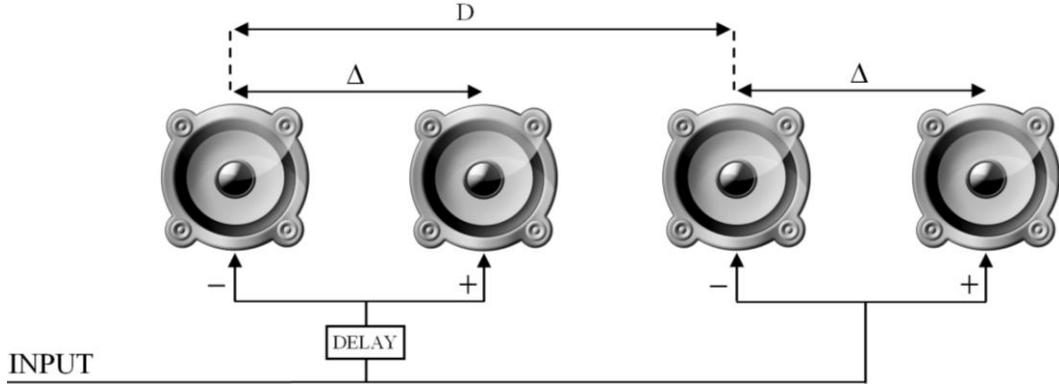


Fig. 4.18 Second-order gradient source configuration

$$R_\theta = \sin\left(\frac{kD}{8} \cos\theta\right) \sin\left(\frac{kD}{4} + \frac{kD}{4} \cos\theta\right) \quad (4.16)$$

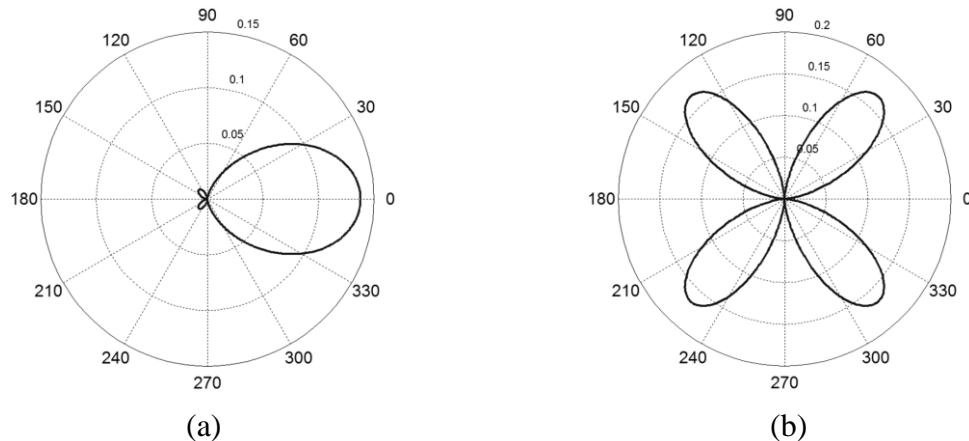


Fig. 4.19 Second-order gradient source polar pattern with drive-unit spacing at
(a) $\frac{1}{4}$ wavelength and (b) full wavelength

Higher order gradient sources are realized by combining zero-, first- and second-order configurations in a similar manner. It is expected that as source order increases, polar pattern becomes increasingly focused. It must be noted, however, that as order increases source efficiency decreases due to destructive interference between drive-units [66]. This is similar to efficiency loss due to additional omnidirectional subwoofers within a room for modal suppression, as discussed earlier.

Modifying a subwoofer's polar pattern not only causes a decrease in efficiency, but can also drastically affect optimal source placement within an acoustic space. A solution to this problem is to determine generalized positions with maximal modal coupling for each common variety of polar pattern (omnidirectional, dipolar, cardioid). As previously discussed, an omnidirectional source's modal coupling is maximized when placed at antinodes and minimized when placed at

nodes. The reverse is true for dipoles since they operate as velocity sources, although dipole coupling can be fine-tuned by adjusting the source's orientation [67].

Cardioid sources have been suggested to be relatively position-independent in terms of room-mode coupling [67]. As with dipolar sources, cardioid sources can be fine-tuned to best suit the acoustical space by adjusting their orientation. In addition to low position sensitivity, cardioid sources have been shown to be unaffected by substantial changes in room absorption and are not greatly influenced by room asymmetry due to their highly directional radiation pattern [5]. Below the first room-mode cardioid sources show no clear advantage over omnidirectional sources and, in fact, can be a poor choice if system efficiency is important.

With this in mind, Backman proposed a subwoofer with a hybrid, frequency-dependent polar pattern [5]. Below the lowest room-mode the subwoofer operates as an omnidirectional source while above that mode it operates as a cardioid source. This approach ensures system efficiency at very low-frequencies while simultaneously reducing source position sensitivity.

Backman highlights two implementations to this system. One approach (Fig. 4.20a) involves applying an all-pass filter (APF) to one drive-unit to modify its phase, achieving the desired frequency-dependent polar pattern. The alternative approach (Fig. 4.20b) places a high-pass filter (HPF) before one drive-unit so only one of the drivers radiates energy below a defined frequency. The first method is preferable since it has the advantage of two drivers radiating below the lowest mode, resulting in better pressurization of the room.

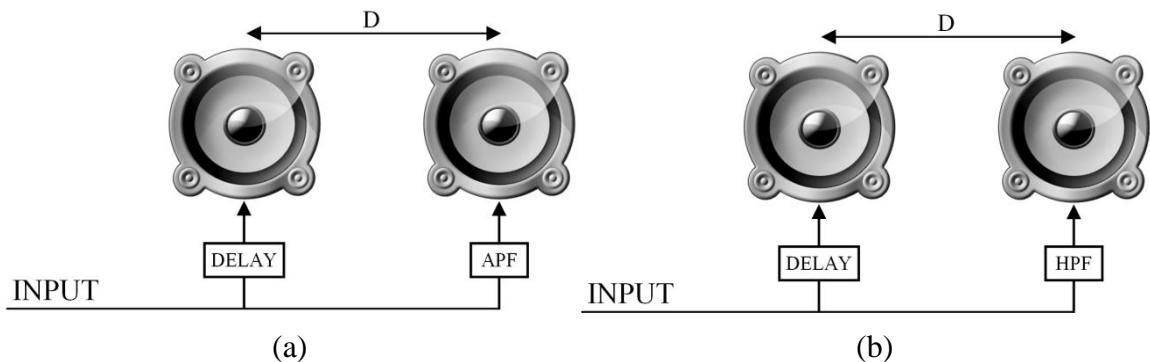


Fig. 4.20 Hybrid loudspeaker polar pattern configurations [5]

Backman's hybrid loudspeaker approach is expanded upon in the following chapter, as its underlying methodology led, in part, to the novel room-mode correction procedure developed over the duration of this research project.

Lastly, it is important to keep the subjective impact of different polar patterns in mind. Linkwitz performs a detailed analysis of perceived sound quality between omnidirectional and dipole subwoofers in [32]. Linkwitz concludes that perceived differences between the polar pattern varieties are largely due to the differing interactions within the room. Omnidirectional sources couple strongly to room-modes and result in longer decay times. Dipole units, on the other hand, were only found to strongly couple to the room-modes along their on-axis direction. This leads to less modal excitation and shorter decay times [32]. This difference between polar patterns must therefore be considered when designing an appropriate low-frequency sound reproduction system.

4.2.5 Absorption

The drawback of passive absorption techniques at low-frequencies is discussed in detail in Section 4.1.2. The fundamental issue is that the required size of the absorbers is on the order of several meters; highly impractical for small-sized listening rooms. The question a handful of researchers have posed, therefore, is whether these passive forms of absorption can be substituted with an active means of absorption.

There are a number of methodologies focused on this question [68 – 71] where most utilize the same core procedure. These systems call for two independently-controllable low-frequency reproduction components: one to serve as the primary source and the other to provide the absorption. The most straightforward implementations [68, 69] focus on a two- or four-source system in rectangular rooms. One/two of the sources are placed on antinodal planes of the low-order axial modes. The secondary source(s) are placed at the opposite wall, mirroring the primary sources. These secondary sources are delayed according to the propagation delay from the primary sources and with their signals inverted and attenuated to compensate for the propagation loss across the room. Ideally, the absorber sources emit signals reflected off the wall (but inverted).

This technique was thoroughly tested in [69], both with simulations and real-world experiments. The system shows a strong ability to suppress modal resonances for the lowest axial room-modes. It does not, however, provide significant correction for higher-order axial or tangential/oblique modes.

This shortcoming was subsequently addressed in [70] where inspiration was drawn from the work in [71]. Instead of using static delay and attenuation to generate the absorption signals, Vanderkooy suggests utilizing microphones placed at or near the absorber sources. This way the signals received at the absorbers can be directly inverted and reproduced from the absorbers to perfectly cancel wall reflections.

This should produce a perfect travelling plane wave along the primary listening dimension. Vanderkooy notes that there is more work necessary for this system to be realized concerning stability, as the measurement microphones need a form of echo cancellation to avoid uncontrollable feedback between the microphones and the active absorbers.

The potential advantages of this active absorption method are clear, nonetheless. Unlike the pure source delay/attenuation in rectangular rooms, this system allows for arbitrary source placement in a space of any non-rectangular topology. At the time of writing, though, a stable active absorption system of this variety has yet to be successfully implemented, although there appears to be a considerable amount of research focused on achieving this goal. There are, however, commercially available systems using similar active absorption routines, such as noise-cancelling headphones.

4.2.6 Radiation resistance

An alternative low-frequency room-mode correction approach to impulse-response measurement-based equalization and active absorption is proposed by Pedersen in [72]. Pedersen's technique involves correction based on radiation resistance measurements close to a loudspeaker. It operates using two measurements. Radiation resistance is first measured using a reference loudspeaker in a reference room. Next, the measurement is repeated in the listening room with the actual loudspeaker. Pedersen claims the relationship between the two measurements can be used to construct an equalization filter that allows the perceived timbre of

the loudspeaker in the listening room to resemble that of the reference loudspeaker in the reference room. This permits a user to choose any target timbre since it is based solely on the reference measurement.

Radiation resistance can be determined by the following method. Given a driver moving at velocity, v , a pressure, p , is created in front of the diaphragm. The pressure on the surface of the diaphragm results in a force, F , acting on the diaphragm. The ratio of force to velocity is referred to as mechanical impedance. In the case of the diaphragm being affected by its acoustical surroundings, this mechanical impedance is referred to as radiation impedance, Z_r (Eq. 4.17) [72].

$$Z_r = \frac{F}{v} \quad (4.17)$$

Radiation impedance is a complex frequency-dependent number. At low frequencies, the imaginary component dominates, which corresponds to a portion of air moving along with the diaphragm. This is equivalent to placing an additional mass on the diaphragm [72]. The real component of the radiation impedance relates to the loss of energy consumed by radiating the acoustic power [72].

Radiation resistance is found using the ratio of sound pressure, p , and volume velocity, q , where volume velocity is found using Eq. 4.18. Another method for calculating radiation resistance (Eq. 4.19) utilizes the radiation impedance found with Eq. 4.17. Radiation resistance has a large influence on the acoustical power output, along with the drive-unit velocity (Eq. 4.20) [72].

$$q = vS_D \quad (4.18)$$

$$R_r = Re(Z_r) = Re\left(\frac{F}{v}\right) \quad (4.19)$$

$$P_a = \frac{1}{2} Re(Z_r)|v|^2 = \frac{1}{2} R_r |v|^2 \quad (4.20)$$

where:	q	=	volume velocity (m/s)
	S_D	=	surface area of diaphragm (m^2)
	R_r	=	radiation resistance
	$Re(Z_r)$	=	real component of radiation impedance
	P_a	=	acoustic power output (W)

Pedersen's proposal utilizes a measurement technique to determine radiation resistance in non-anechoic spaces. This involves taking two measurements of the sound pressure near the surface of the diaphragm (close, but separate locations). The radiation resistance is then calculated based on the relationship between the two pressure measurements (Eq. 4.21) [72]. Pedersen asserts that the transfer function of the microphone has very little influence on the radiation resistance measurement.

$$R_r = Re\left(jgf \frac{p_1}{p_1 - p_2}\right) \quad (4.21)$$

where:	g	=	scaling factor (frequency independent)
	p_1	=	sound pressure at measurement point 1
	p_2	=	sound pressure at measurement point 2

The radiation resistance measurements must be performed both in the reference room and the listening room. Once the measurements are obtained, the low-frequency correction filter is generated by the relationship given in Eq. 4.22 [72].

$$EQ(f) = \sqrt{\frac{R_{r,ref}(f)}{R_{r,lst}(f)}} \quad (4.22)$$

where:

$EQ(f)$	=	complex filter coefficients
$R_{r,ref}$	=	reference radiation resistance measurement
$R_{r,lst}$	=	list. room radiation resistance measurement

The velocity of the diaphragm will be affected by applying this filter, as shown in Eq. 4.23. This ensures the acoustic power output in the listening room is identical to that in the reference room (with the reference loudspeaker) (Proof 4.1) [72].

$$v_{lst}(f) = EQ(f)v_{ref}(f) \quad (4.23)$$

where:

$v_{lst}(f)$	=	listening room diaphragm velocity (m/s)
$v_{ref}(f)$	=	reference room diaphragm velocity (m/s)

$$P_{a,ref} = \frac{1}{2} R_{r,ref} |v_{ref}|^2$$

$$P_{a,lst} = \frac{1}{2} R_{r,lst} |v_{lst}|^2$$

$$P_{a,lst} = \frac{1}{2} R_{r,lst} \left| v_{ref} \sqrt{\frac{R_{r,ref}}{R_{r,lst}}} \right|^2$$

$$P_{a,lst} = \frac{1}{2} R_{r,lst} \frac{R_{r,ref}}{R_{r,lst}} |v_{ref}|^2$$

$$P_{a,lst} = \frac{1}{2} R_{r,ref} |v_{ref}|^2 = P_{a,ref} \quad (\text{Proof 4.1})$$

The benefit of this correction method is that it does not require measurements at multiple listening locations, but gives roughly the same results. The filters are dominated by boundary and room-mode effects, but are essentially the same as impulse-response measurement-based multiple point filtering routines [72]. It is unclear how the method will function when placed at nodal points for one or more room-modes, where source-to-room coupling is theoretically zero. This may result in the system essentially ignoring those few modes, which gives an incomplete solution to the problem.

4.2.7 Ambisonics-style equalization

The active low-frequency room correction procedure highlighted in this section differs from the standard impulse-response measurement design by being based on a fundamental principle of ambisonics [73]. The method assumes a sound field can be completely characterized by four

parameters, consisting of an omnidirectional channel (W) and three dipole channels (X, Y, Z) oriented along each of the primary dimensional axes.

In this method, the sources must be setup surrounding a single listener. Like in most conventional systems, the impulse response is measured at the listening location for each source component in turn. These measurements are then combined to create an equalization filter for each source component giving the target response at the single listening location (Fig. 4.21) [73].

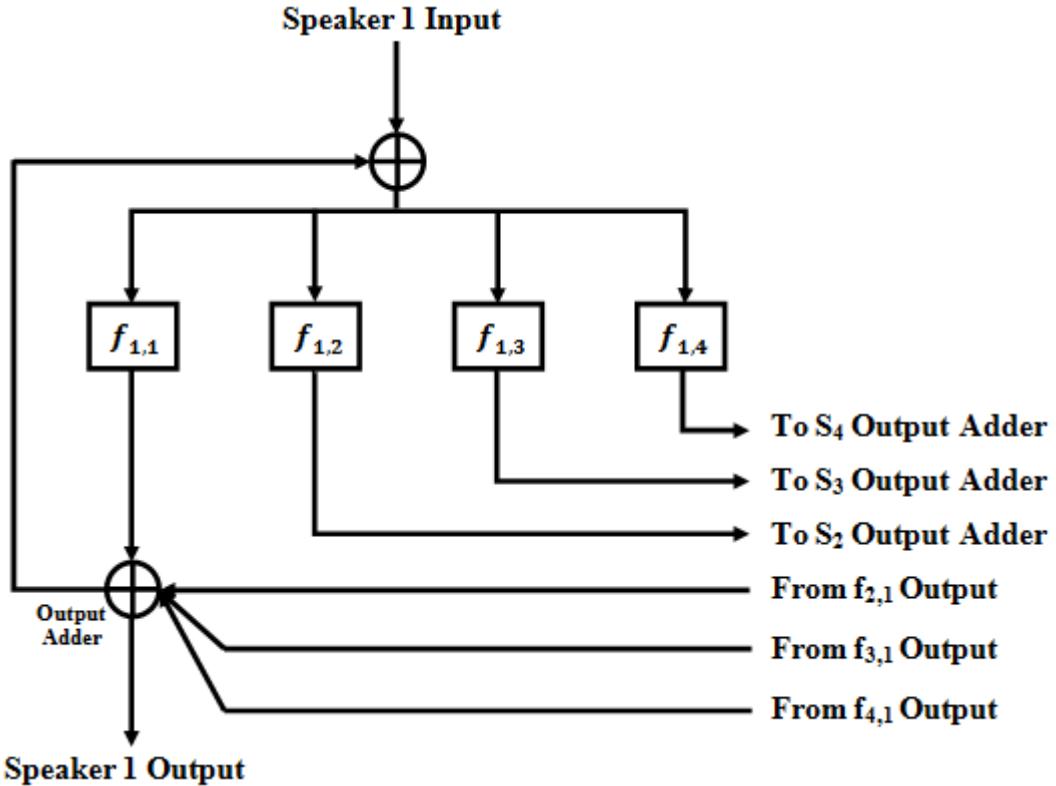


Fig. 4.21 Ambisonics-style room correction (only one source component shown) [73]

This method could possibly be expanded to include multiple listening locations and also adjusted to handle arbitrary source placement. This possibility serves as inspiration for aspects of this research project, which is discussed in detail in Chapter 5.

4.2.8 Active correction conclusions

A wide-variety of active low-frequency correction techniques have been described, highlighting the strengths and weaknesses of each. The bulk of these methodologies rely on impulse-response measurements at one or more location across a listening area. Some of these procedures operate statically, with a one-time filter generation, while others operate by continuously updating the filters (requiring measurement microphones in the listening area at all times), although both function using similar error-signal minimization routines. The most common minimization technique is the least-mean squares method, although arguments have been made for other approaches.

Multiple-point equalization approaches are shown in published research to provide acceptable spatial variance reduction, by grouping similar responses in the error minimization function and weighting them accordingly. None of the described research in this area mentions how non-

measurement points in the vicinity of the target points benefit (or not) from equalization. Also, these methods require what can be a time-consuming optimization procedure that could make real-time system adjustment difficult.

Polar pattern control, active absorption, radiation resistance correction and ambisonics-based correction are alternative approaches. These methods have been (and continue to be) perused by significantly fewer researchers, but each shows promise of achieving high levels of spatial variance reduction. A combination of these techniques has served as an impetus for the work to be described in Chapter 5.

4.3 Chapter summary

There exist numerous methodologies concerning low-frequency spatial-variance control in small-rooms. These methods can be categorized into two varieties: passive and active, although some approaches utilize concepts spanning both varieties.

Passive correction generally involves the physical characteristics of a closed space along with the layout of a sound reproduction system. These approaches have proven beneficial for controlling the modal spectral distribution to avoid sharp resonances and also to suppress strong low-order axial modes by means of low room-to-source coupling or destructive interference. Few of these methods, however, significantly contribute to spatial variance reduction. The multiple-subwoofer positioning approach is an outlier to this statement.

Active correction methods aim to correct for what passive approaches cannot. Most active techniques involve one-time or continuous measurements in the listening area in order to generate one or more equalization filters aimed to create a highly accurate spatiotemporal low-frequency response. These methods can require high computational resources and aren't likely to lend themselves easily to real-time adjustments. That point aside, these approaches have taken large strides towards achieving the stated low-frequency reproduction goals.

The various successes of passive and active correction approaches have been taken into account in this research, with the goal of developing a system that embraces the advantages of both varieties, including the non-mainstream active methods such as polar pattern control and an ambisonics-like approach for wide area correction. The relevant methods are touched upon again in the remaining chapters.

5 Chameleon subwoofer arrays

The existing procedures and proposals regarding low-frequency room-mode correction generally assume subwoofer units are of the omnidirectional, single degree of freedom variety (highlighted as zero-order gradient sources in Section 4.2.4). This restriction limits low-frequency correction as most practical sound reproduction systems only accommodate up to four subwoofers (although most entry-level commercially-available systems only support a single subwoofer). Theoretically, each system degree of freedom (single drive-unit subwoofer) allows for correction at a unique listening location. More degrees of freedom allow for independent control over more locations.

There is, however, a handful of work that alludes to the possibility of using a single subwoofer with multiple independently-controllable drive-units or multiple omnidirectional subwoofers with individual processing for each [5, 19, 66, 73]. In [66], Olson presents methods for low-frequency polar pattern control using multiple drive-units, with physical separation, electronic delay and/or polarity reversals applied to each driver, as needed (as discussed in Section 4.2.4). This approach, while frequency-independent, provides the first argument towards multiple degrees of freedom in loudspeakers to better control the room response, although low-frequency modal problems were not the focus of his work.

Backman expands on Olson's gradient loudspeaker techniques, motivated specifically by the low-frequency room-mode problems encountered in small-sized listening rooms [5]. As discussed in Chapter 4, Backman notes that although dipolar subwoofers are useful in suppressing modal excitation, they rapidly lose efficiency as frequency decreases; a point also emphasized by Olson [66]. In order to maintain system efficiency to the lower limit of the reproduced frequency range, Backman devised a two drive-unit subwoofer scheme generating a quasi frequency-dependent polar pattern using either an all-pass filter or a high pass filter on one of the drive-units [5]. This results in a dipolar coverage pattern over the higher subwoofer band, but transitions to an omnidirectional pattern below a specified frequency to boost efficiency. Backman notes that below the lowest room-mode spatial variance is no longer an issue as the subwoofer simply pressurizes the room, where no standing waves can occur [5].

While Olson and Backman focus on the issue of polar pattern control by multiple drive-unit subwoofer designs, Howe and Hawksford concentrate less on the specific characteristics of the loudspeaker and more on the overall effect of the sound reproduction system [73]. Their approach is rooted in ambisonics where a listener is surrounded by loudspeakers, both horizontally and vertically. The encompassing loudspeaker system allows for the delivery of a three-dimensional sound field with accurate control over localization effects. Instead of concentrating on the localization utilities of the ambisonics methodology, Howe and Hawksford approach the system as a means for equalizing the room response, thus eliminating any unwanted acoustical artifacts of the room, such as strong room-modes [73]. The system was demonstrated to provide accurate correction over an 80 cm diameter sphere while targeting a 2 m x 2 m “region of equalization” around the target listening location (up to 200 Hz).

Regarding room-mode correction over multiple listening points, Welti proposes a system which operates using multiple omnidirectional subwoofers which are individually controllable to achieve minimal spatial variation between listening locations [19]. Each subwoofer is provided a uniquely controllable FIR filter for correction purposes, with the filter coefficients found using an optimization routine. This system was demonstrated to drastically reduce spatial variance in the subwoofer band, although the optimization routine often takes several minutes to complete and the results are limited by the number of single drive-unit subwoofers that can

be accommodated within the listening space [19]. Welti notes that a direct calculation procedure is possible provided there are an equal number of degrees of freedom in the subwoofer system and listening locations [19].

The above mentioned work provides motivation towards the work for this project. The system presented in [73] lends a very powerful tool for room equalization over a wide listening area, although the method requires that the user be surrounded by sources. The methods in [5] and [66] demonstrate that polar pattern control is achievable using multiple drive-units within a subwoofer and frequency-dependent patterns can be achieved with simple filters [5]. The system presented in [19] offers a powerful solution to the room-mode problem, however optimization can be time-consuming and utilizing many single drive-unit subwoofers can be impractical in constricted environments.

This chapter focuses on the development and implementation of what are referred to as *chameleon subwoofer arrays* (CSA). The theory behind this new system is rooted in the work in [5, 19, 66, 73], but granting greater flexibility concerning the source layout while also providing fully frequency-dependent polar pattern control, where each frequency bin can ideally have a different coverage pattern to best suit the needs of the acoustical space.

This chapter begins with an in-depth presentation on the theoretical concept behind CSA technology, including various system stability issues. This is followed by analysis of simulation results using the bespoke FDTD simulation toolbox. Analysis includes examination of correction benefits both at, and away from, the target measurement points, as well as the relationship between the correction efficiency and the physical system layout. CSA correction results are directly compared to the most common correction procedures in order to determine the advantages and disadvantages of the new practice.

5.1 Theoretical concept

The chameleon subwoofer array (CSA) correction system takes into account the highlighted shortcomings of existing solutions to low-frequency room-mode problems (as discussed in Chapter 2). The relevant components of this concept are described in detail in the following subsections, highlighting key differences between CSAs and existing approaches.

5.1.1 Degrees of freedom

Sound reproduction manipulation flexibility is directly related to the number of degrees of freedom present in a system. One single drive-unit subwoofer only provides correction at a single point, as is highlighted in the discussion on single-point equalization in Chapter 4. An additional subwoofer theoretically allows for correction at a secondary point. Following this line of reasoning, each single drive-unit subwoofer contributes a degree of freedom to the correction system, thus allowing for another listening point response to be directly manipulated. This assumes the processing and amplification stages can accommodate the additional independent drive-signals.

Welti demonstrates in [19] that more degrees of freedom in a reproduction system increases control precision over a listening area. This is closely related to the ambisonics approach presented by Howe and Hawksford [73] where each system component is individually controlled to provide maximum control over the response at the listening location. The CSA system, therefore, must deliver as many degrees of freedom as possible while holding to practical physical limitations. Toole suggests in [2] that up to four subwoofers can be practically placed in a conventional listening room. More than four subwoofers consume too

much space, not to mention that it is doubtful any consumer would invest in such a large number of units. With this in mind, a four subwoofer limit can be placed on CSAs (although larger arrays will be explored for completeness).

5.1.2 Hybrid sources

Hybrid, multiple drive-unit, subwoofers are necessary to provide more degrees of freedom while keeping the number of subwoofers to four or less. Backman's proposal of a two drive-unit subwoofer [5] is a step towards what is necessary. Backman's system, however, has each subwoofer operating without regard to other system components. Ideally, system components should adjust themselves so that all active components collaboratively achieve the desired coverage pattern.

CSAs are designed to operate with one or more hybrid subwoofer. These hybrid units adopt the wisdom of ambisonics where spherical harmonics of various orders are utilized to create complex coverage patterns [73]. In this case, the hybrid subwoofers consist of six drive-units; one on each side of a sealed cubic enclosure. This configuration allows for four primary spherical harmonic components; one zero-order component (omnidirectional), where all drive-units are driven synchronously, and three first-order components (dipolar) each oriented in one of the three spatial dimensions.

The first-order spherical harmonics (dipoles) are realized by driving the components on opposing sides of the enclosure with one unit having reverse polarity, as per Olson's gradient loudspeaker methodology [66]. Theoretically, this will produce a perfect figure-of-eight pattern along each primary dimension, however may be difficult to achieve in practice. As the hybrid subwoofer is a sealed enclosure, there may exist a level of source component cross-talk, thus blurring the resulting polar pattern. Additionally, cabinet diffraction effects (largely manifested as shifts in the acoustical center of the individual drive units [21, 22]) can cause further difficulty in achieving the figure-of-eight pattern since these first-order gradient sources are sensitive to drive-unit spacing, as highlighted by Olson in [66].

It can be argued that providing independent control to each drive-unit gives superior control over combining the four spherical harmonic components. Upon careful inspection, though, it is clear that addressing the drive-units individually is not advantageous as each unit contributes a roughly omnidirectional radiation pattern (at low-frequencies, relative to the enclosure dimensions) and since the drive-units are positioned in close proximity to one another it can be surmised that the room responses from each drive-unit will exhibit high correlation. The set of spherical harmonics, on the other hand, are orthogonal and can be combined to produce a specific polar pattern.

5.1.3 Mathematical structure

The CSA correction procedure operates using impulse response measurements at each target listening location with one source component activated at a time. These measurements are taken using an MLS, generally of the 12th or 13th order at a sampling rate around 4 kHz (as the frequency band of interest extends to around 200 Hz). This length MLS allows the room to reach steady-state so an accurate assessment can be made of the acoustical characteristics.

The MLS measurements are processed to determine the impulse response at the target locations due to each activated source component (see Chapter 2 for details). The impulse response measurements are converted to the frequency domain via a fast Fourier transform (FFT) to allow for correction filter coefficient generation. If there are more listening locations than

source components, the superfluous locations are ignored. The same adjustment is performed if there are more source components than listening locations. The system operates assuming an exact match in number between listening locations and source components (unless pseudo-inverse filtering is utilized, as detailed in Section 5.5.3).

The frequency responses are placed into an $NL \times NS \times NF$ matrix where NL , NS and NF are the number of listening locations, source components and frequency bins, respectively. This matrix is labeled FR and is analyzed one frequency bin at a time. To simplify the calculation procedure, a matrix labeled X is loaded with the $NL \times NS$ data matrix corresponding to the current frequency bin under inspection.

A target response must be determined as the CSA correction system steps through the frequency bin range. This can be any response, but Vanderkooy argues that a good target response for low-frequency correction is the room average response (which is rarely used in commercial products, which generally target a flat response) [70]. Vanderkooy asserts that humans are accustomed to listening to the acoustical characteristics of a space and it would be nonsensical to remove these characteristics. Therefore, the target response for each listening location is determined at each frequency bin by taking the average response over all measurements. This assumes a linear system, where superposition of the individual source component responses is valid. While nonlinearities will certainly exist in practice, the linearity assumption should be sufficiently accurate to allow for effective correction. This claim is supported by a similar assumption for the correction procedure developed in [19] which exhibits many parallels to the procedure described in this research and has been shown to provide accurate low-frequency correction in a wide-range of listening environments. Target responses need not be identical, but for this initial exploration of CSA correction, they are set to give identical objective responses at each location. Individualized target responses are discussed in Chapter 7.

Listening location target responses are arranged into an $NL \times 1$ matrix labeled Y . This matrix operates at one frequency bin at a time for memory conservation purposes. The unknowns in the CSA are the correction filter coefficients, contained in an $NS \times 1$ matrix labeled H . Assuming system linearity, the process of target response calculation is shown in Fig. 5.1.

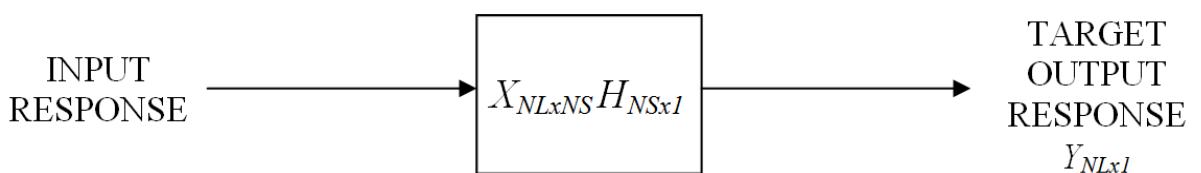


Fig. 5.1 Basic CSA correction procedure block diagram

In order to calculate the correction coefficients, the X matrix (containing the uncorrected frequency responses) can be inverted and multiplied by the target response matrix, Y . This gives the correction coefficient matrix, H (Eq. 5.1). This calculation requires an equal number of source components and measurement points (a square X matrix), however this restriction can be subverted using pseudo-inverse filtering which is explored in Chapter 5.5.

$$H_{NSx1} = X_{NLxNS}^{-1} Y_{NLx1} \quad (5.1)$$

Once filter coefficients are calculated for each frequency bin they must be transformed back into the time domain so they can be applied to the input signal via convolution. This is

accomplished with an inverse fast Fourier transform (IFFT). Since correction is ignored above the Nyquist frequency, all frequency bins above said limit are initially zero. To generate a purely real-valued digital filter, these frequency bins must be set to the reflection of the complex conjugate of the first half of the frequency range (excluding the DC term) over the Nyquist frequency bin.

Since each filter impulse response is likely to exhibit slight pre-ringing, a circular shift must be applied so that the peak of the response lies at the center of the vector. Once the filters are converted to the time domain they are convolved with the source signal to generate individual drive signals for the source components.

Of course, care must be taken to ensure practical solutions, in terms of filter stability and coefficient amplitude. This is addressed in the following sections.

5.2 System practicality issues

As mentioned in Section 5.1.3 and in [19], a direct calculation method for correction filters can potentially cause unrealistic requirements on the sound reproduction system. These problems arise for a number of reasons including: over-excursion of dipole pairs, attempted correction outside the discrete modal range and also problematic source-to-listener coupling where correction at certain mode(s) requires an unrealistically large amount of acoustical energy to address one or more areas of a listening space.

5.2.1 Dipole efficiency

The work in [32] and [66] emphasizes that dipolar loudspeakers lose efficiency as frequency decreases due to increased cone excursion requirements to overcome the destructive interference between the two drive-units used to create the directional pattern. This efficiency loss is approximately 6 dB per octave for first-order gradient loudspeakers [66]. Considering these performance losses at very low frequencies it is clear how problems arise with what is essentially blind correction, unconcerned with the physical limitations of the drive-units.

This problem requires examination of the targeted listening area. The upper region of the subwoofer frequency band (~45 Hz and above) contains multiple wavelengths of each frequency within the listening area. In this range there is low correlation between adjacent points' responses since they lie sufficiently apart within the waveforms of the frequencies.

In the lower region of the subwoofer frequency band (below ~45 Hz) the measurement points yield data that exhibits very high correlation since points are relatively close to each other. This limits precision correction as the measurement points are strongly correlated (and naturally exhibit minimal spatial variance between one another) which can limit the invertibility of the measurement matrix [75]. This can result in high-amplitude correction coefficients required to achieve only minimal spatial variance reduction. In an attempt to correct for such cases it is likely that the dipolar units will face unrealistic cone-excursion requirements, which should be avoided.

The solution is to split the correction frequency range into two bands. As the listening area is oversampled in the lower band, some measurement points and source components are unnecessary. The dipolar source components should be inactive in this frequency range to avoid over-excitation issues and measurement points only along the periphery of the listening area will be utilized. If, for example, there are four hybrid subwoofers (containing sixteen source components), then in the lower subwoofer band only the four omnidirectional components will

be used with the four corresponding measurement points. All sixteen source components and measurement points become active in the upper subwoofer band.

Calculation of the crossover point, f_x , (Eq. 5.2) is based on the width of listening grid (in measurement points), N_m , the width of the room (m), l_x , the spacing ratio of the listening points (1.5, 2 or 3 from the FDTD toolbox), sp , the omnidirectional source component measurement point spacing ratio, oR , and the mean point spacing (m), H_{sp} (Eq. 5.3).

$$f_x = \frac{c}{4H_{sp}oR} \quad (5.2)$$

$$H_{sp} = \frac{l_x}{spN_m} \quad (5.3)$$

Equation 5.2 states that the crossover point, f_x , is inversely proportional to four times the product of the mean measurement point spacing with the omnidirectional measurement point spacing ratio. The spacing ratio represents how many high-band (dipole + omnidirectional) points exist between each omnidirectional point. Generally, this ratio is 2 or 3, depending on the size of the measurement grid. The product of this ratio with the average point spacing effectively gives the mean omnidirectional measurement point spacing which is critical in defining the crossover frequency (where the crossover is set to the frequency which one-quarter wavelength fits perfectly between two omnidirectional points).

Equation 5.3 determines the mean measurement point spacing by calculating the width of the grid (by multiplying the number of points along one side of the grid with the user-defined point spacing ratio, sp , which spaces the points based on the width of the room). The width of the room is then divided by the calculated width of the grid to give the mean measurement point spacing.

The crossover frequency ensures that only frequencies where at least a quarter-wavelength fits within the listening area are corrected for using the full omnidirectional and dipolar source component system. When less than a quarter-wavelength fits in the listening area only the omnidirectional source components are used to ensure system efficiency and to prevent dipole unit overload. The crossover point is not used as an abrupt transition; rather there is a smooth transition between bands with an overlap of $\pm 5\%$ surrounding the crossover frequency. There is an additional factor concerning dipole efficiency that relates to the physical placement of the dipole source components in relation to the room boundaries where this topic is explored in Section 5.5.

5.2.2 Lower correction limit

The acoustical characteristics of a space can also cause problems within a direct calculation correction procedure. As discussed in Chapter 2, sound pressure spatial variance over a listening area is directly related to room-modes. Below the lowest mode, however, spatial variance largely disappears since wavelengths at these frequencies do not fit within the space. Since the sound reproduction system is pressurizing the room in this range, there should only be minor variations in response over the listening area. These minimal variations result in tightly coupled listening location responses which can be difficult to correct for independently due to the wavelength dimensions being on the order of the primary room dimensions. Attempting to precisely control this range of frequencies could require unrealistic demands on the system,

only to achieve minute reductions in spatial variance (where only a small variance existed in the first place).

Along with the relatively minor spatial variance of sound pressure at very low frequencies, it is also helpful to note the human ear's sensitivity to pressure changes in this range. Fletcher and Munson show in [74] that as frequency decreases, the perceived level change is generally less than the objective change in level. For instance, at 1 kHz a reduction in SPL from 80 to 70 dB is perceived as a 10 dB drop, but at 30 Hz that same 10 dB reduction is perceived as an attenuation of less than 2 dB. The underlying point is that at very low-frequencies the human ear has poor sensitivity to pressure level changes. Modest spatial variation in this range is likely to be difficult to perceive. Therefore, precise correction is less critical in this band.

The lower correction limit, f_L , is calculated using the mean spacing of the omnidirectional measurement points, L_{sp} , as a guide (Eqs. 5.4 & 5.5). The cutoff frequency also depends on the size of the listening area. A smaller listening area results in a higher cutoff frequency while a larger area lowers the cutoff frequency. In a sixteen measurement point configuration, for example, the cutoff point is set at the lowest frequency where a sixth of a wavelength fits within the listening area. Below that limit, correction is not performed.

$$f_L = \frac{c}{3L_{sp}oR} \quad (5.4)$$

$$L_{sp} = H_{sp}oR \quad (5.5)$$

5.2.3 Upper correction limit

The boundary between the discrete modal (low-frequency) and diffuse (high-frequency) bands in small rooms was discussed in Section 2.2. This boundary is referred to as the Schroeder frequency. Above this frequency room-modes are sufficiently dense in the spatial and spectral domains so that they are not distinctly noticeable to the human ear. Therefore, above the Schroeder frequency modal correction is not necessary. The response in this band can be corrected, if necessary, using the non-subwoofer loudspeakers in the system.

Spatial sampling must also be considered. Similar to time-domain sampling where the upper frequency limit for accurate sampling is defined by the frequency which receives only two sample points over its wavelength, spatial sampling requires at least two sample points in space to give accurate sampling. In this particular instance, the sampling points are the measurement locations dispersed throughout the listening area (generally in a grid layout). The highest frequency expected to benefit from CSA correction, f_H , is therefore related to the mean spacing between adjacent measurement points (Eq. 5.6).

$$f_H = \frac{c}{2H_{sp}} \quad (5.6)$$

Although Eq. 5.6 gives the absolute upper limit for the correction system, this value is likely to be above the subwoofer band. To ensure correction targeted only within the operational range of the subwoofers, the upper correction limit is reduced to the subwoofer crossover point if Eq. 5.6 gives a higher cutoff frequency.

5.2.4 Windowing

Aside from unrealistically large filter coefficient amplitudes, filter ringing can cause pre- and post-echoes. A practical solution is to apply a gentle window to the correction filter impulse responses to remove extraneous ringing. The shape and length of the window must be chosen carefully, as it is crucial to preserve as much of the correction data in the filter as possible. A window that is too short will remove most correction benefits. An overly wide window will be unable to address the worst-case ringing. The goal is to choose a window that can handle all cases, from the best (no ringing) to the worst (heavy persistent ringing).

A range of windows were tested using eight source/listener configurations (Table 5.1). All configurations were placed in a virtual 5 m x 4 m x 3 m room with 20 cm grid-point spacing and 10% surface absorption. Correction was applied from 20 – 120 Hz based on impulse response measurements using a 13th order MLS. Four hybrid subwoofers were utilized along with a 16-point listening grid, thus granting sixteen degrees of freedom. All correction results were judged using a 36-point walking path through the listening area. The benefits of using a walking path are discussed in Section 5.4.

In order to judge the effectiveness of CSA correction using various windows, the uncorrected system spatial variance over a walking path is calculated for each system configuration (Table 5.2). The correction results are judged based on the percent reduction in spatial variance from the uncorrected to the corrected systems.

#	Subwoofer #1 (m)	Subwoofer #2 (m)	Subwoofer #3 (m)	Subwoofer #4 (m)	Listening Grid ζ (m)	Description
1	(0.6, 0.6, 0.6)	(4.4, 0.6, 0.6)	(0.6, 3.4, 0.6)	(4.4, 3.4, 0.6)	(2.2, 1.8, 1.8)	Room corner placement
2	(0.6, 0.5, 0.6)	(0.6, 1.5, 0.6)	(0.6, 2.5, 0.6)	(0.6, 3.5, 0.6)	(3.0, 1.8, 1.8)	Room front placement
3	(1.2, 0.6, 0.8)	(1.2, 0.6, 2.2)	(1.2, 3.4, 0.8)	(1.2, 3.4, 0.8)	(3.0, 1.8, 1.8)	2 x vertical stacks (1/4 room width placement)
4	(0.6, 1.0, 0.6)	(0.8, 2.2, 0.6)	(1.4, 3.2, 0.6)	(1.6, 0.8, 0.6)	(3.0, 1.8, 1.8)	Random placement
5	(0.6, 2.0, 0.6)	(4.4, 2.0, 0.6)	(2.5, 0.6, 0.6)	(2.5, 3.4, 0.6)	(2.2, 1.8, 1.8)	Wall midpoints (ground)
6	(0.6, 2.0, 1.5)	(4.4, 2.0, 1.5)	(2.5, 0.6, 1.5)	(2.5, 3.4, 1.5)	(2.2, 1.8, 1.8)	Wall midpoints
7	(0.6, 0.6, 0.6)	(4.4, 0.6, 2.4)	(0.6, 3.4, 0.6)	(4.4, 3.4, 2.4)	(2.2, 1.8, 1.8)	Room corner placement (rear units on ceiling)
8	(4.4, 1.8, 0.6)	(1.0, 1.6, 0.6)	(3.8, 3.4, 0.6)	(0.6, 3.2, 0.6)	(2.2, 1.8, 1.8)	Random placement

Table 5.1 Room configurations for CSA window choice evaluation

#	Spatial variance (dB)
1	3.13250
2	5.37180
3	4.98610
4	4.09160
5	4.33410
6	4.26280
7	3.22250
8	4.53930

Table 5.2 Uncorrected spatial variance of various system configurations detailed in Table 5.1

The windows evaluated are designed to cover the entire time frame of the correction filter impulse response in order to truncate the response at the extremes where ringing should be suppressed. The reduction in spatial variance between the uncorrected and corrected systems due to each window was determined using simulations and is displayed in Fig. 5.2.

The windows tested are:

- Rectangular (no window)
- Blackman-Harris
- Chebychev (80 dB side lobe attenuation)
- Gaussian ($\alpha = 2.5$)
- Hamming
- Hanning
- Kaiser ($\beta = 3.0$)
- Triangular

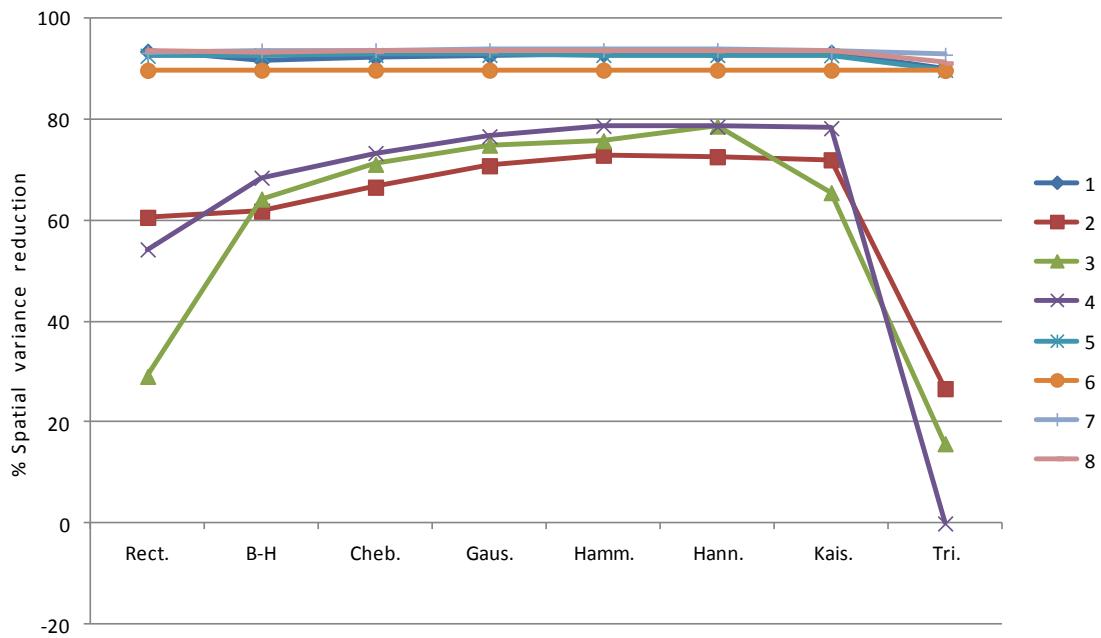


Fig. 5.2 Spatial variance reduction between uncorrected and CSA corrected systems due to various window selections

This window evaluation provides insight into the operation of the CSA. There are clearly configurations that benefit more from CSA correction than do others. Configurations 1, 5, 6, 7 and 8 each exhibit 90% or more spatial variance reduction. Correction for these configurations appears to be largely independent of window choice, indicating that the correction filters are stable with only modest ringing (Fig. 5.3). Configurations 2, 3, and 4, on the other hand, show a high dependence on window choice. In these three cases, the triangular window gave the worst results with configuration 4 not benefiting at all from correction.

At first glance, it may seem odd that the rectangular window isn't the best performing since it doesn't alter the correction filters in any way, which should give an ideal solution to the problem. Considering, however, that correction performance was judged over a virtual walking path (where most points do not coincide with the measurement points), it is understandable how the rectangular windows can perform poorly (Fig. 5.4).

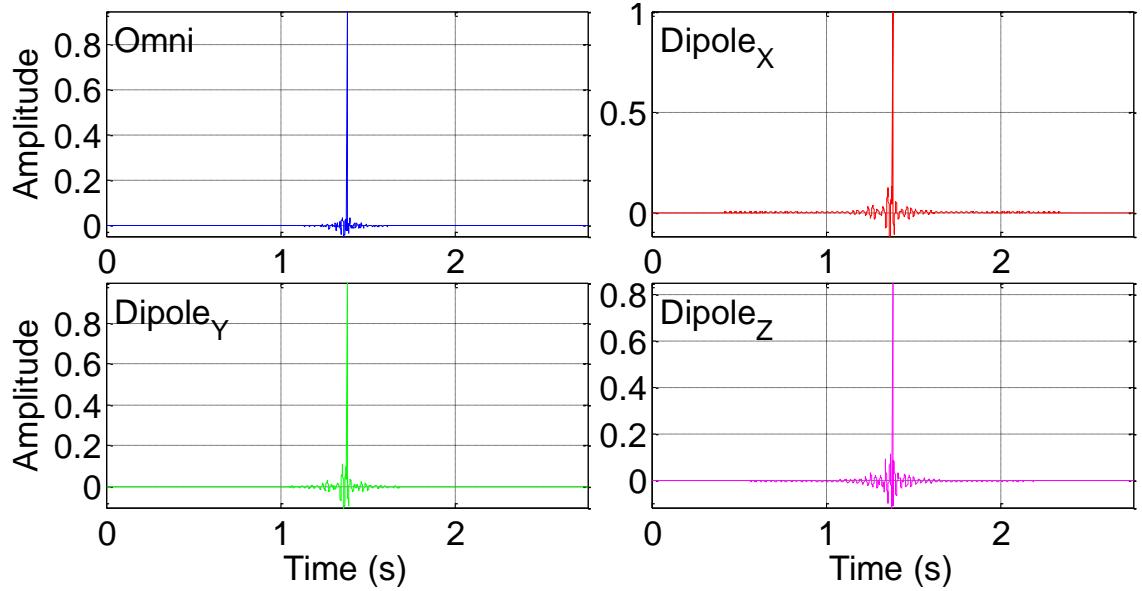


Fig. 5.3 CSA correction filter impulse responses exhibiting minimal ringing

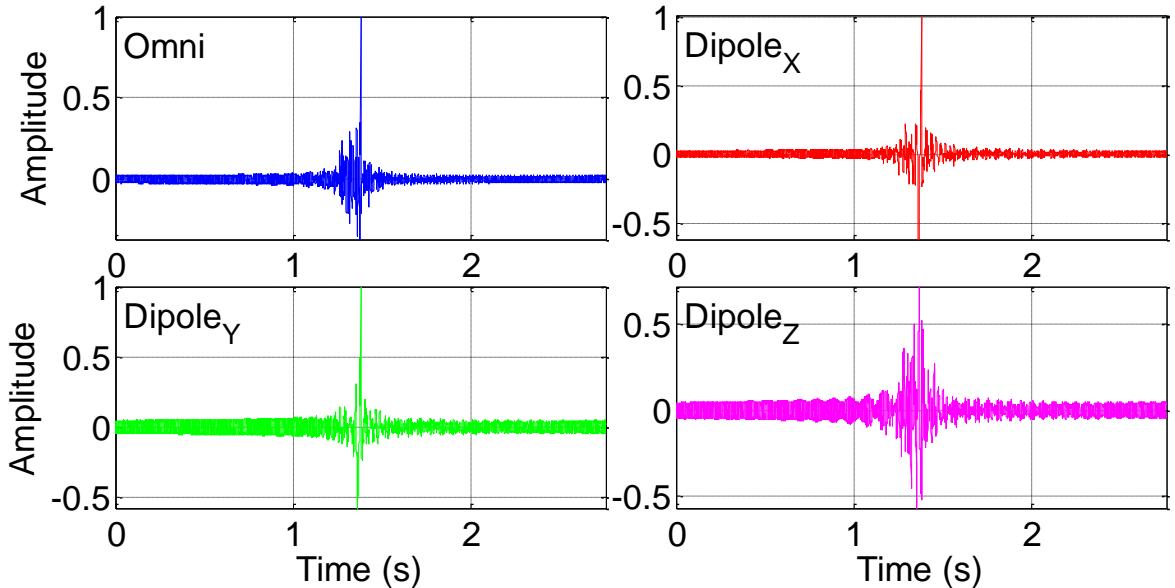


Fig. 5.4 CSA correction filter impulse responses exhibiting persistent ringing

The best performing windows for the three worst configurations are the Hamming and Hanning windows, giving at least 70% spatial variance reduction. Since the Gaussian and Kaiser windows allow for adjustment of the width of the window in MATLAB, tests were conducted to examine how the widths of these two window types affect the overall correction benefits and if a certain width choice outperforms the Hamming/Hanning windows. The α and β values for the Gaussian and Kaiser windows were adjusted from 0.0 (no window) to 128 (extremely narrow window) and then inspected for any trends (Figs. 5.5 & 5.6).

There are clear trends in the data obtained from the Gaussian and Kaiser window width tests. Once the window width reduces to the point where a significant amount of data is lost, correction benefits drop sharply and can actually worsen the response over the listening area in extreme cases. Additionally, if the window is too wide ringing may not be entirely suppressed, introducing errors into the corrected system.

Based on the presented results, a Kaiser window with $\beta = 6.0$ is chosen as it gives consistent results over all tested configurations where most configurations show prime performance around $\beta = 4.0$ with an estimated peak in performance falling slightly above this point (based on the data in Fig. 5.6). All CSA correction results presented from this point onwards utilize this window, ensuring immunity to system instabilities based on this and the bandwidth limitation procedures.

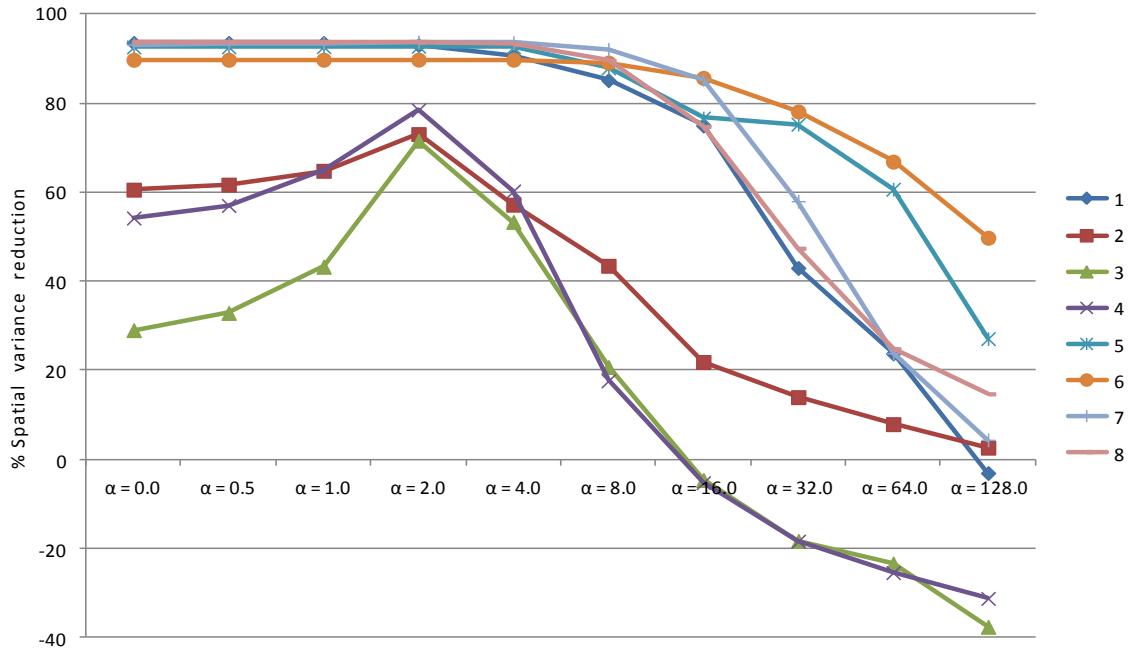


Fig. 5.5 Spatial variance reduction between uncorrected and CSA corrected systems due to various Gaussian window α values

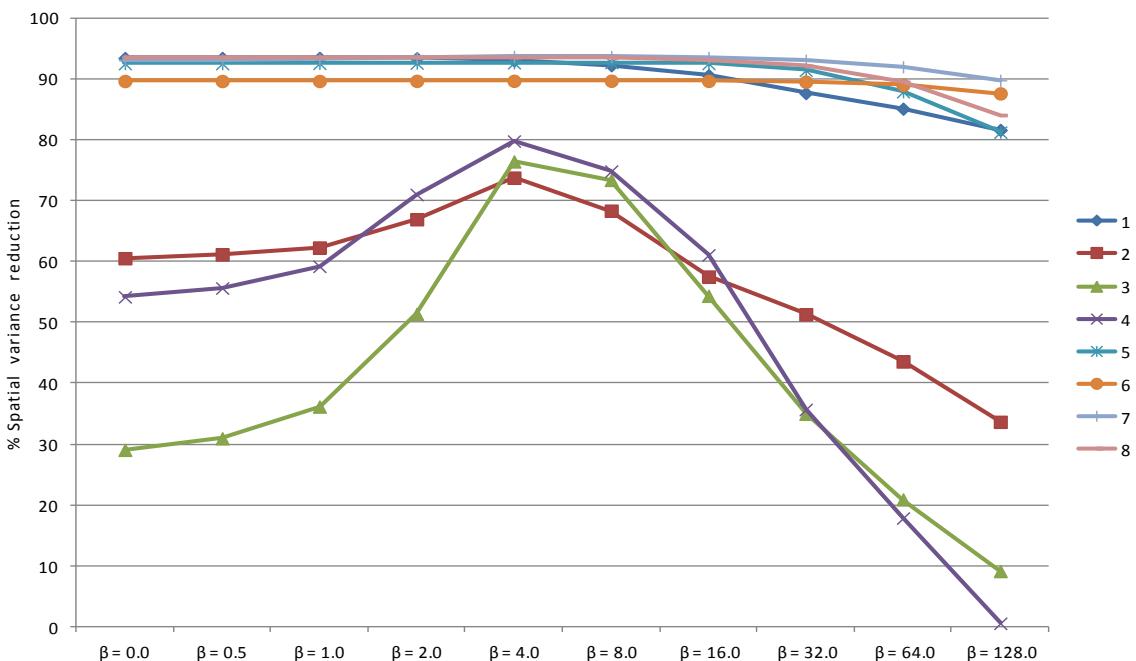


Fig. 5.6 Spatial variance reduction between uncorrected and CSA corrected systems due to various Kaiser window β values

5.3 Correction performance

Having addressed CSA correction system consistency in Section 5.2, focus can be placed on the performance of the system over a wide area. Before an entire listening area is considered, the target measurements points must be examined to judge how they benefit from the CSA algorithm. Since measurements from these points are used directly to generate the correction filters, they should show close agreement to the target responses after correction.

A three-dimensional virtual space was set up in the FDTD toolbox of dimensions 8 m x 6 m x 3.5 m with grid point spacing of 20 cm and 5% surface absorption. Three system configurations were utilized for these tests: a one-, two- and four-unit CSA with the corresponding number of listening locations. Focus was placed both on correction accuracy and also the corrected system's efficiency (uncorrected versus corrected system mean output level comparison) to determine if and how the target points benefit from CSA correction.

5.3.1 One-unit CSA (four degrees of freedom)

The first configuration utilizes a one-unit CSA, giving four degrees of freedom, necessitating four measurement points. The single hybrid subwoofer (S_1) is placed in one of the room corners at one meter from each surface with the measurement grid placed approximately at the center of the room at a height of 2.4 m (Fig. 5.7).

Before correction, a series of simulations are run to judge the uncorrected system performance. These simulations include an impulse response measurement to determine the steady-state frequency responses (Fig. 5.8) which allows for the calculation of the sound pressure spatial variance between points. In addition to the impulse response measurements, two tone-burst simulations at 35 Hz and 80 Hz (15 cycles per burst, three repetitions) illuminate the transient characteristics over the listening area (Fig. 5.9).

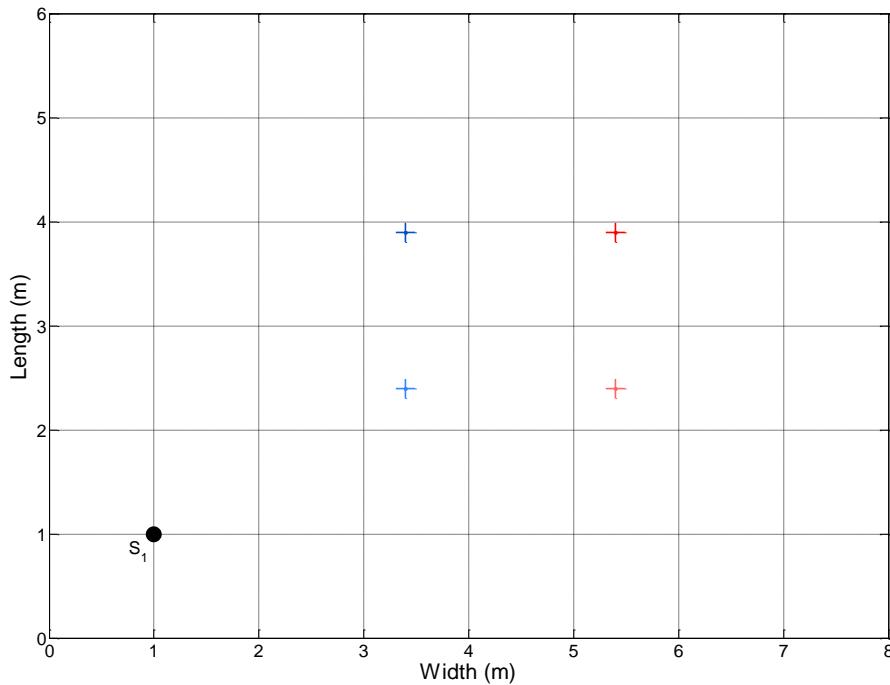


Fig. 5.7 One-unit CSA test system configuration
(Note: target point colors are used in Figs. 5.8 & 5.9)

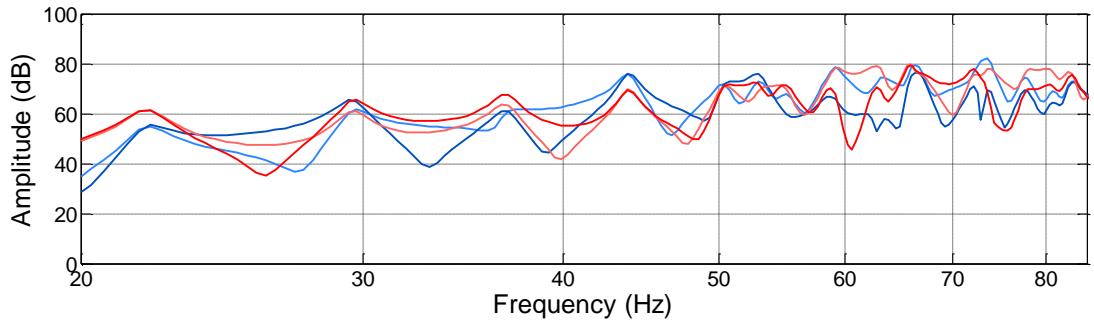


Fig. 5.8 Simulated target point frequency responses for the uncorrected system

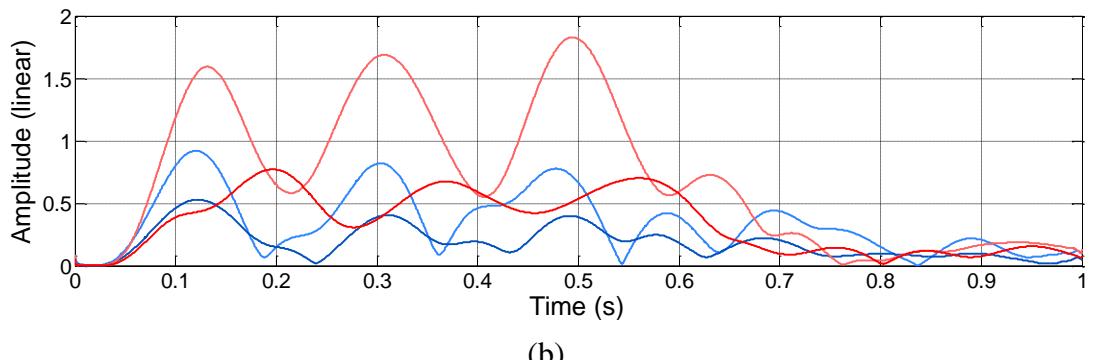
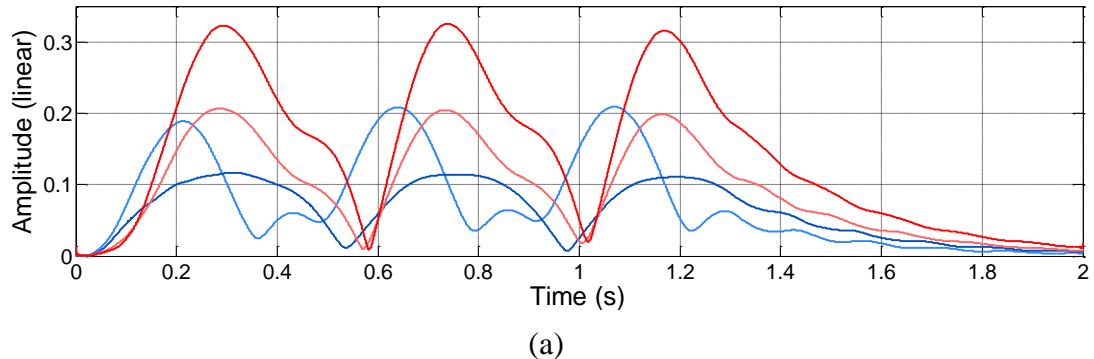


Fig. 5.9 Uncorrected system tone burst simulations at (a) 35 Hz and (b) 80 Hz

The uncorrected system highlights the fundamental issues of wide-area low-frequency sound reproduction. In Fig. 5.8 there are significant deviations between each target point frequency response, resulting in an overall spatial variance of 5.12 dB (20 – 85 Hz, the reason for the frequency range is discussed later). Listeners dispersed throughout the room receive different low-frequency responses, making it difficult to reliably deliver adequate listening experiences to everyone.

The tone burst tests in Fig. 5.9 highlight the transient response issues the uncorrected system exhibits. In both trials, the three bursts arrive at the target points in very different manners. Some points clearly distinguish the three bursts, while at the same time adjacent points do not due to transient smearing as well as reduced amplitude. Both of these issues must be addressed by CSA correction.

Next, the system has CSA processing applied to create an even spatio-pressure distribution over the listening area. Due to the wide spacing of the target points, CSA correction can only be applied up to 85 Hz (based on the limit imposed in Eq. 5.6). The correction filters' complex coefficients and impulse responses are shown in Fig. 5.10.

The correction filters enforce the concepts discussed regarding dipole loudspeakers, both in this chapter and in Section 4.2.4. The unit is placed near a room corner; therefore its position is coincident with the antinodes of most room-modes. The omnidirectional component, which operates as a pressure source, strongly couples to these room-modes resulting in a high output. The dipole units, on the other hand, operate as velocity sources and weakly couple to room-modes when placed at antinodal positions. This requires more energy to drive the dipole units to produce the required output.

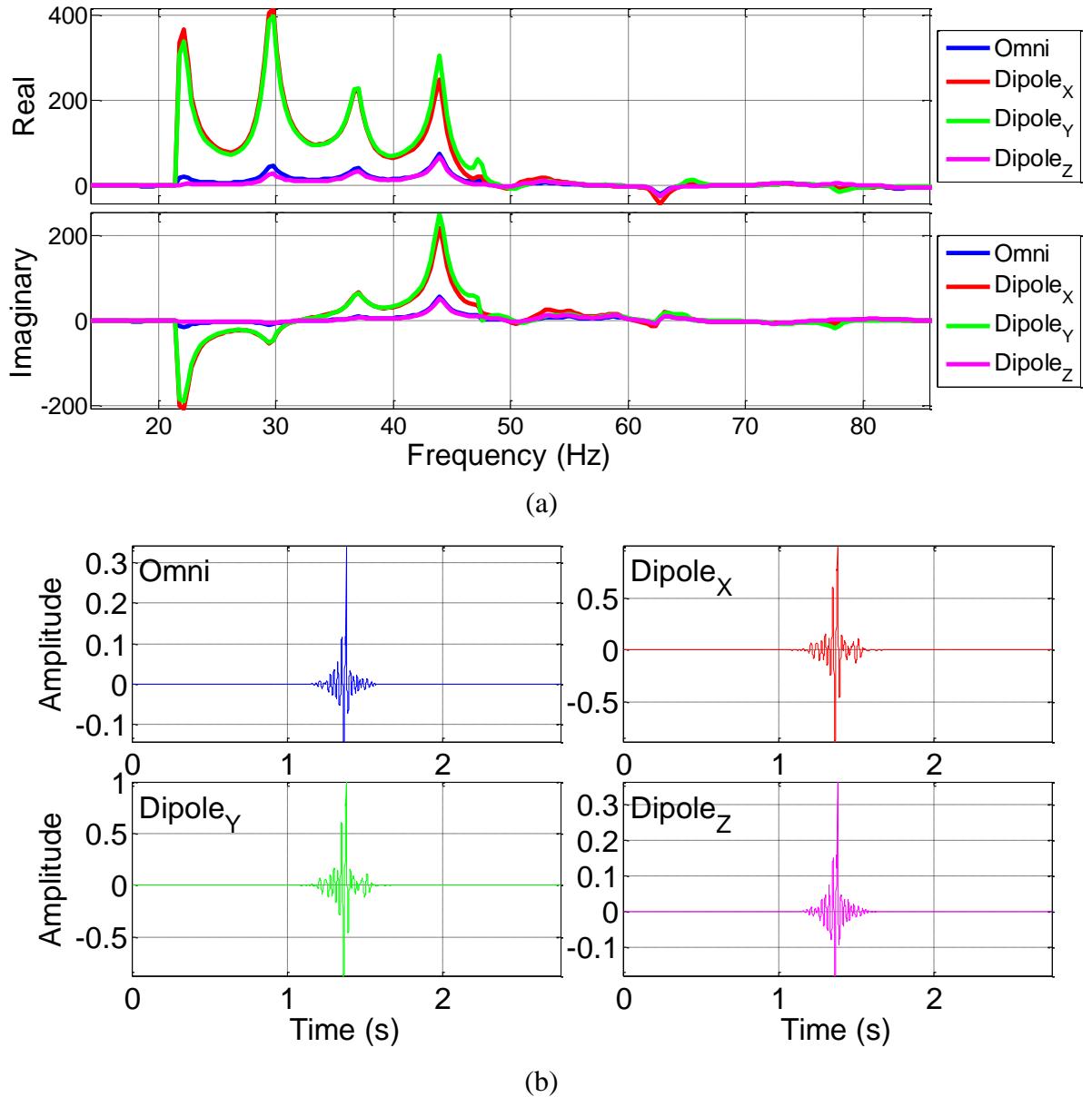


Fig. 5.10 CSA correction filter (a) complex coefficients and (b) impulse responses

The increased output requirements of the dipoles are evident in the correction filter coefficients for the x- and y-dipole components in Fig. 5.10. The z-dipole component does not have as high correction requirements since all measurement points lie in a horizontal plane, resulting in very low measured spatial variation at the vertical room-modes. These heavy requirements on two of the dipole units lower system efficiency, as compared to the uncorrected system. This comparison is accomplished by repeating the impulse response and tone bursts simulations used to judge the uncorrected system (Figs. 5.11 & 5.12).

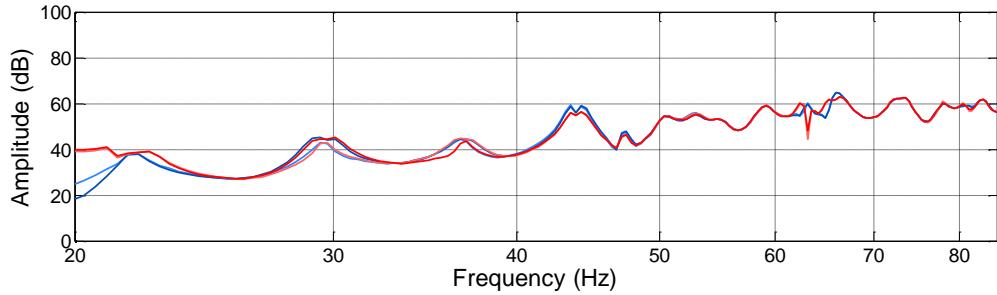


Fig. 5.11 Simulated target point frequency responses for the CSA corrected system

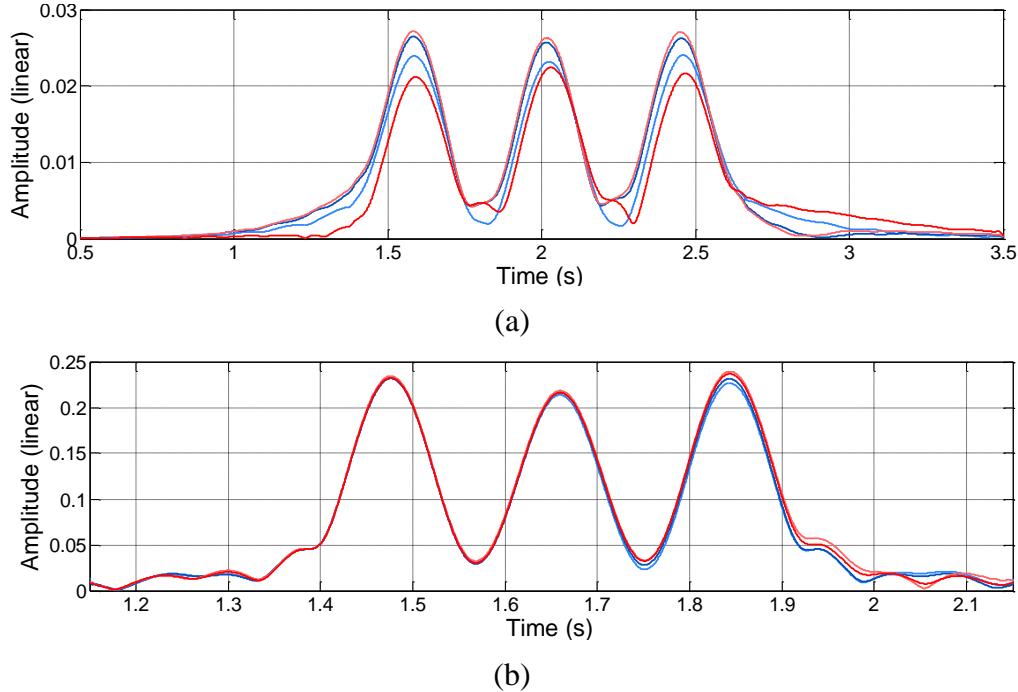


Fig. 5.12 CSA corrected system tone burst simulations at (a) 35 Hz and (b) 80 Hz

Compared to the uncorrected system, CSA correction reduces spatial variance from 5.120 dB to 0.785 dB, a reduction of 84.7%. Similarly, the tone burst tests show a marked improvement as each listening location receives an objectively clear reproduction of three pulses (Fig. 5.12). It must be noted that the corrected tone burst measurements are not perfectly in line with the source signal since the CSA correction targets the room average response as opposed to a linear, uncolored (flat) response.

While the frequency band from 50 – 85 Hz sits around 80 dB SPL for the uncorrected system, it drops to 60 dB SPL after correction. The 20 dB output loss is a consequence of a protective feature in the algorithm to limit cone excursion of the drive-units. Once the impulse responses are generated, they are normalized so that the maximum amplitude over the filter set is one. Fig. 5.10a highlights the excessive correction requirements for the x- and y-dipole units due to corner placement. These large coefficients reduce the amplitude of the generated impulse responses, which gives a lower overall system output. It is assumed that the subwoofer is sufficiently raised above the floor so that the z-dipole operates efficiently.

The hybrid subwoofer can be shifted away from the room corner to increase the effectiveness of the dipole units. This is accomplished as shown in Fig. 5.13 with the correction filters in Fig. 5.14 and the uncorrected and corrected target point frequency responses in Fig. 5.15.

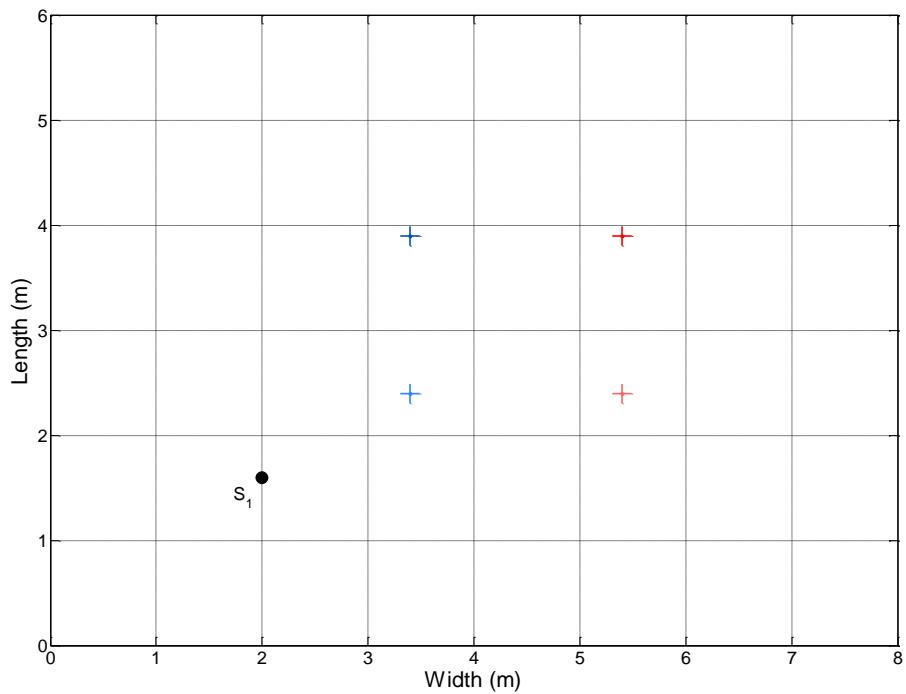


Fig. 5.13 One-unit CSA test system configuration

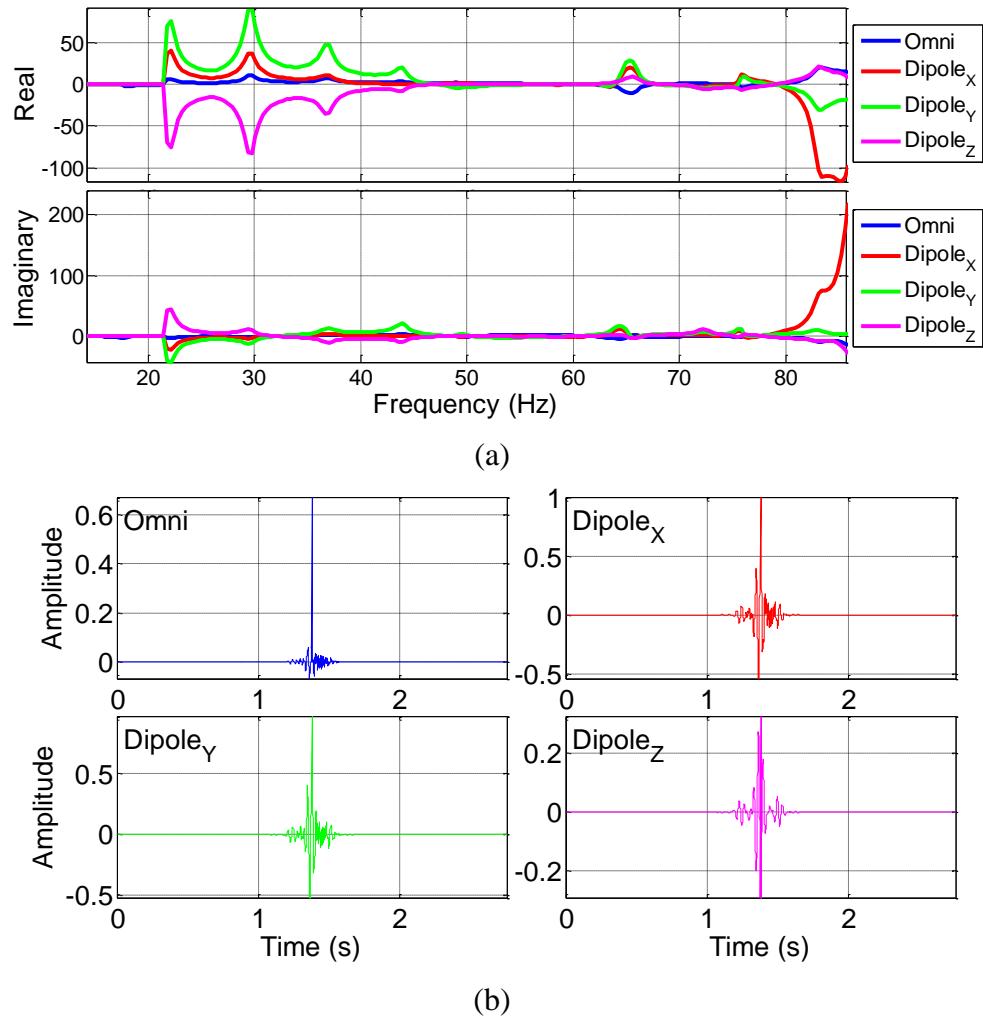


Fig. 5.14 CSA correction filter (a) complex coefficients and (b) impulse responses

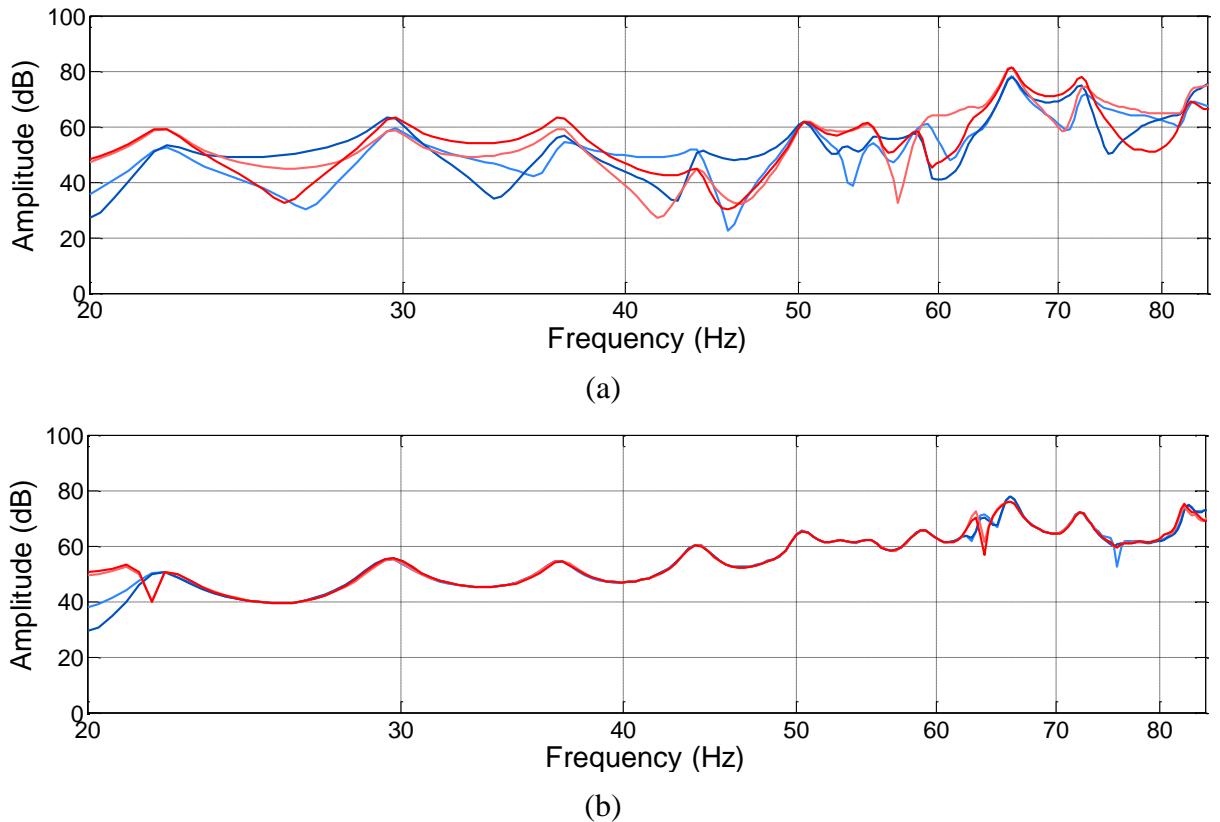


Fig. 5.15 Simulated target point frequency responses for the (a) uncorrected and (b) CSA corrected systems

The effectiveness of the dipole components has been restored by moving the subwoofer less than a meter farther away from the room corner. This is seen both in the more practical correction coefficients (Fig. 5.14) and in the uncorrected/corrected system comparison in Fig. 5.15. In this configuration, the SPL levels are approximately the same before and after correction, indicating no efficiency loss. Spatial variance reduction improves due to the better operation of the dipoles, now exceeding 88%.

This exploration highlights the importance of source placement relative to room boundaries, as dipole units become quite limited when near a surface. Moving units further away from boundaries increases system efficiency, albeit while sacrificing the benefits of the Waterhouse effect, as detailed in Section 2.3.2.

Lastly, it is critical that CSA correction benefits all points over a listening area; not just the target measurement points. This is explored with a virtual walking path through the area encompassed by the target points, just as an individual would wander through a space in the real world (Fig. 5.16). The resulting frequency responses are shown in Fig. 5.17.

The walking path simulation highlights the issue with loosely spaced target points. When a listener moves away from these points (spaced by 2 m in this case) the benefits of the CSA correction are potentially lost. This is evident in Fig. 5.17b where between 35 – 45 Hz there exists significant spatial variation, largely stemming from the points along the walking path that stray furthest from the target points (near the center of the listening area). Due to these errors there is only a 39.2% reduction in spatial variance which is an improvement, nonetheless, but it cannot be accepted as a robust solution to the problem at hand.

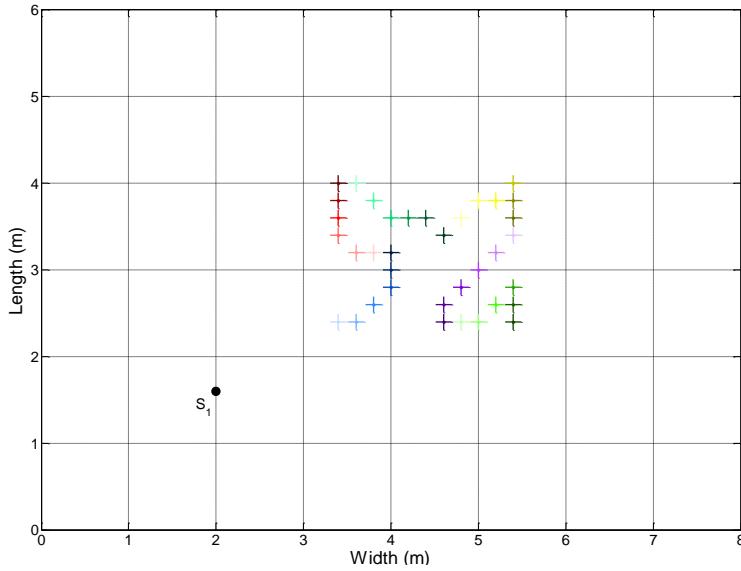


Fig. 5.16 One-unit CSA test system configuration with 36-point walking path

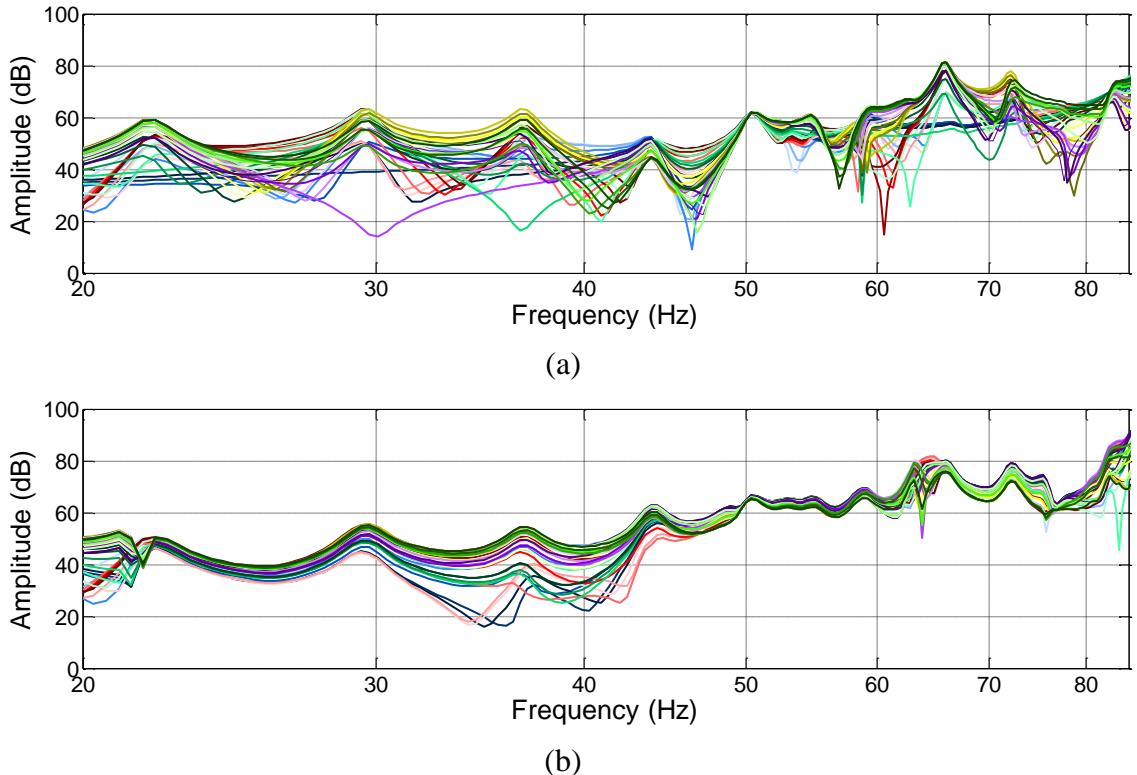


Fig. 5.17 Simulated frequency responses for the (a) uncorrected and (b) one-unit CSA corrected systems over a 36-point walking path

5.3.2 Four-unit CSA (sixteen degrees of freedom)

The practical limit of four subwoofers in a closed space has been described in Sections 4.1.4 and 5.1.1. Bearing this in mind, the system configuration can be expanded to this limit by introducing three additional hybrid subwoofers in the remaining room corners (Fig. 5.18). Again, the filter coefficients and corresponding impulse responses can be calculated with the

CSA algorithm (Figs. 5.19 & 5.20). The additional subwoofers should ease system requirements as they naturally exhibit a lower spatial variance as shown in Table 5.2.

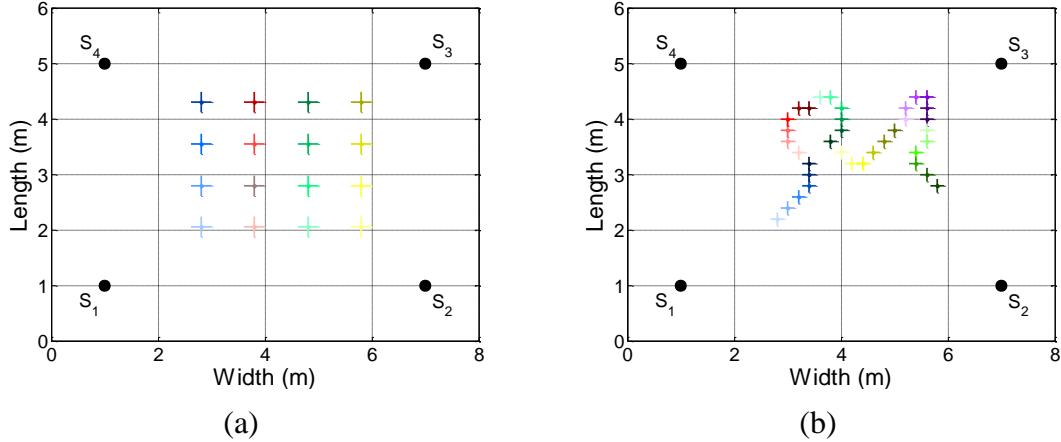


Fig. 5.18 Four-unit CSA configuration for (a) calibration and (b) 36-point walking path

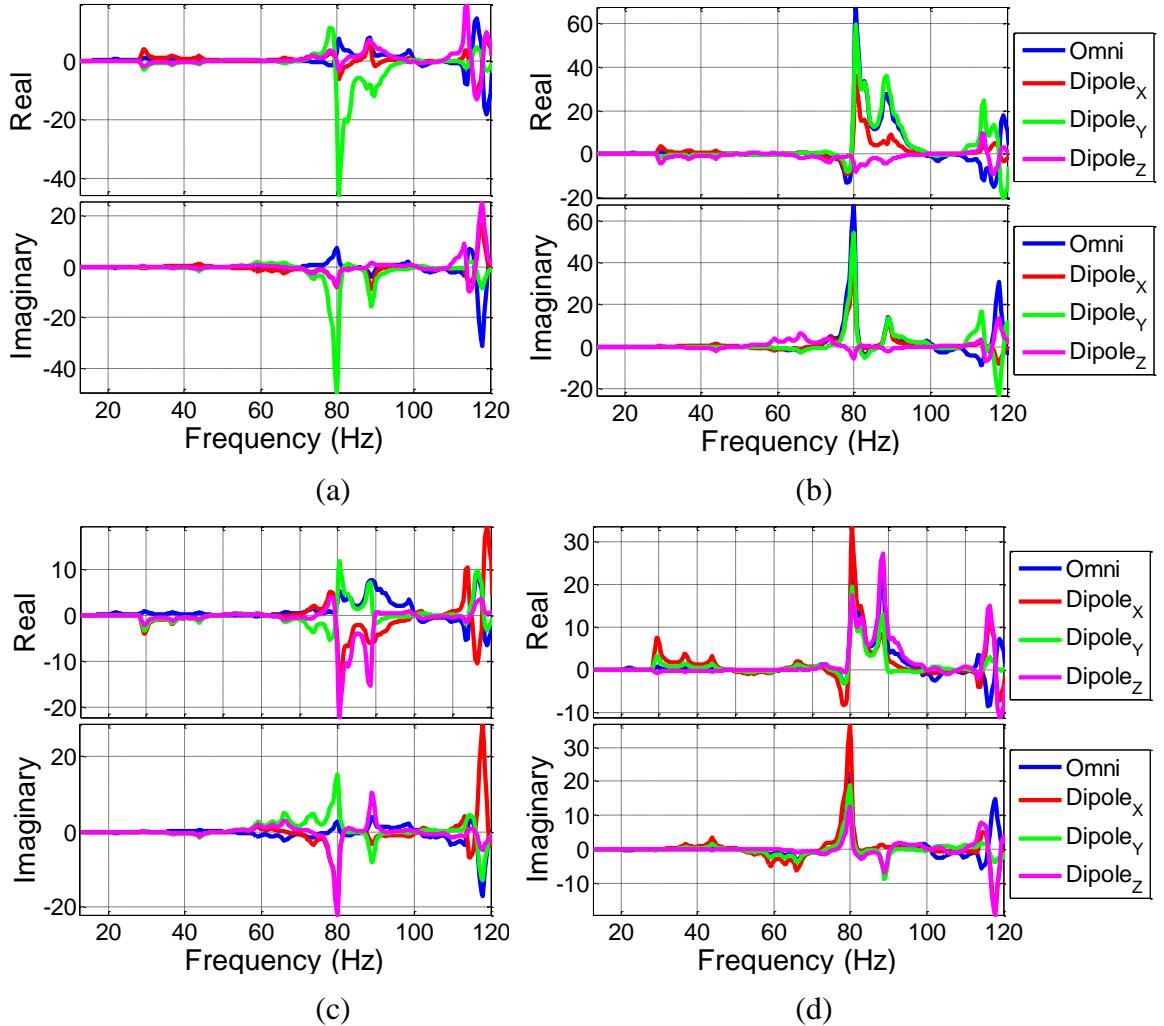


Fig. 5.19 CSA correction filter complex coefficients for (a) hybrid source #1, (b) hybrid source #2, (c) hybrid source #3 and (d) hybrid source #4

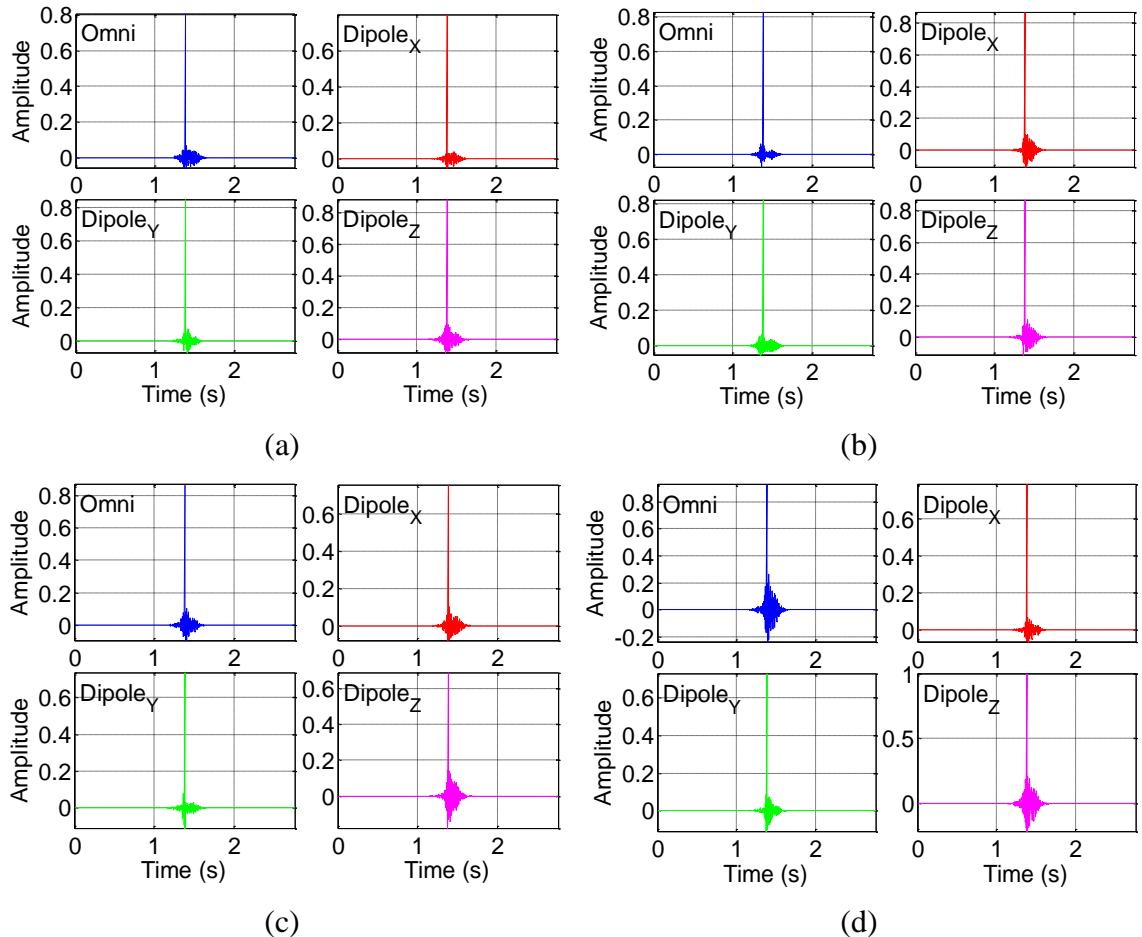


Fig. 5.20 CSA correction filter impulse responses for (a) hybrid source #1, (b) hybrid source #2, (c) hybrid source #3 and (d) hybrid source #4

The correction coefficients in Fig. 5.19 confirm that the positioning of the additional subwoofers has contributed to the efficiency of the correction system. The frequency response and tone burst tests were repeated to judge the correction benefits along the walking path using the four-unit CSA with subwoofer corner placement (Figs. 5.21 & 5.22).

The four-unit CSA gives greater correction benefits than the single-unit system. The walking path benefits from a 91.2% spatial variance reduction (97.3% at target points). The frequency responses over the walking path (Fig. 5.21) are virtually identical, with the largest variance occurring around 80 Hz, although this variance is only around 3 dB. The same holds for the tone-burst responses in Fig. 5.22. Each listening point receives three distinct pulses, with only minor amplitude variation for the 80 Hz case, which is consistent with the steady-state response in Fig. 5.21.

The four-unit CSA does exhibit a loss in efficiency after correction (approximately 10-15 dB lower output). This can be attributed to one or two causes. First, as seen in the one-unit CSA, corner placement isn't ideal for dipoles. This causes a loss of efficiency, especially if the correction filters depend heavily on the dipole components. Secondly, the four-corner configuration results in high destructive modal interference, which suppresses many low-order room-modes. Since the CSA attempts to reinsert these modes into the corrected response, this is essentially working against the system's natural modal suppression.

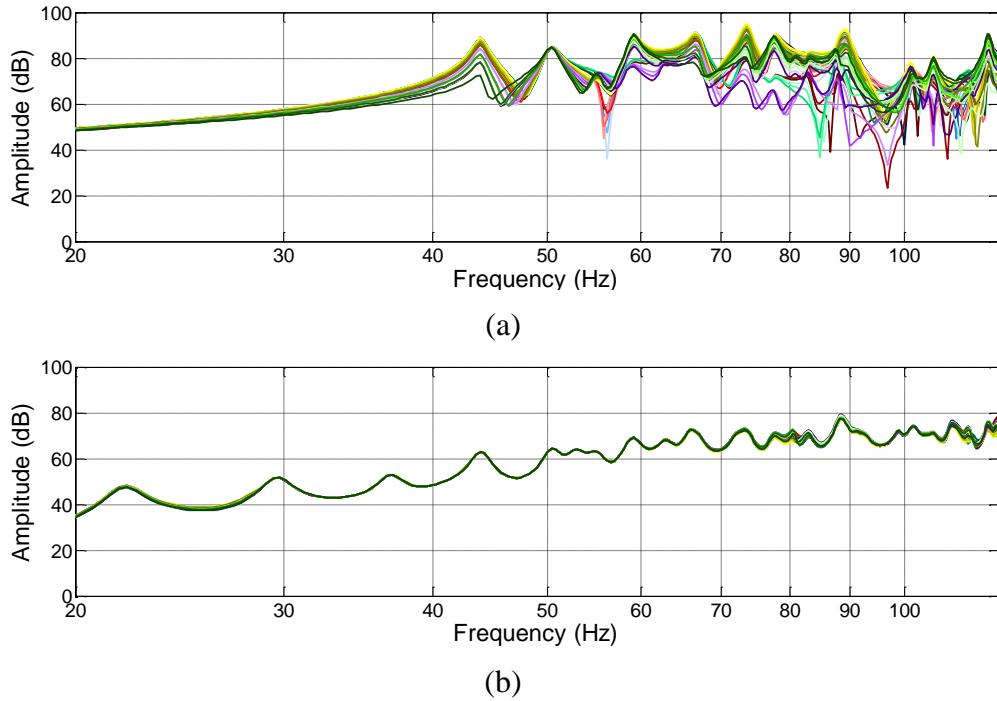


Fig. 5.21 Simulated frequency responses for the (a) uncorrected and (b) four-unit CSA corrected system over a 36-point walking path

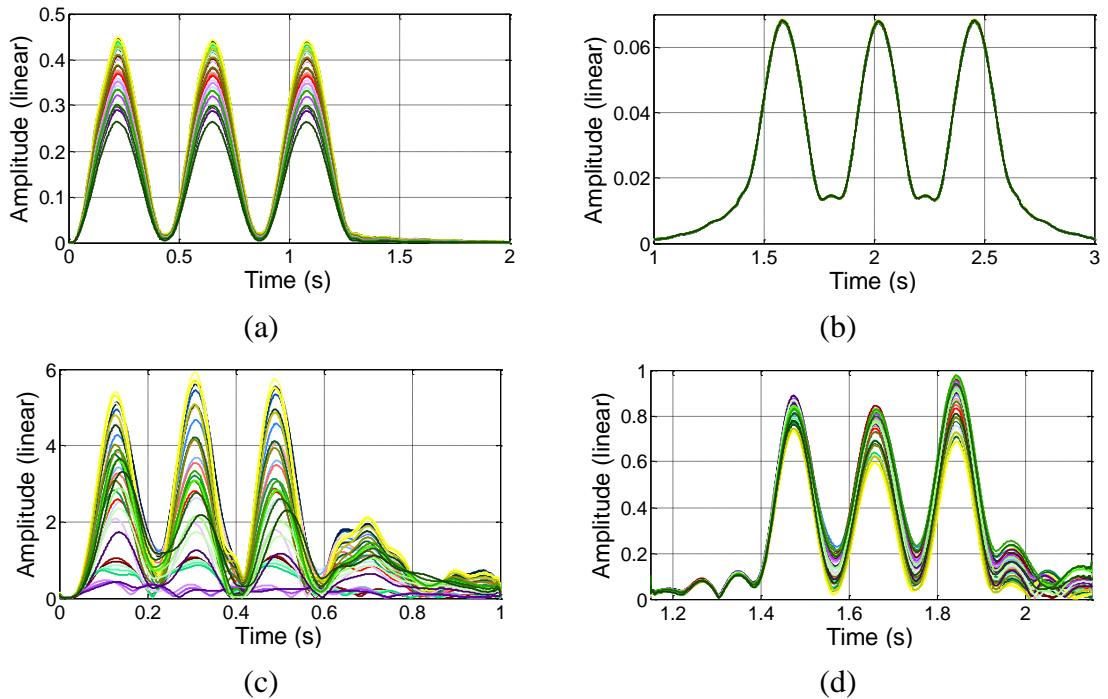


Fig. 5.22 Tone burst responses for the uncorrected system at (a) 35 Hz and (c) 80 Hz and for the corrected system over a 36-point walking path at (b) 35 Hz and (d) 80 Hz

5.3.3 Vertical correction range

Equally important to providing correction benefits over a horizontally-configured listening area is simultaneously providing the same benefits over a vertical range. The four-unit CSA was maintained for this experiment and frequency responses were determined in the same manner, but with the listening grid offset vertically by plus or minus 0.2, 0.4, 0.6 and 0.8 m in individual trials.

As vertical offset increases spatial variance also increases, beginning at higher frequencies due to their shorter wavelengths. The overall trend of spatial variance reduction as a function of vertical listening grid offset is displayed graphically in Fig. 5.23.

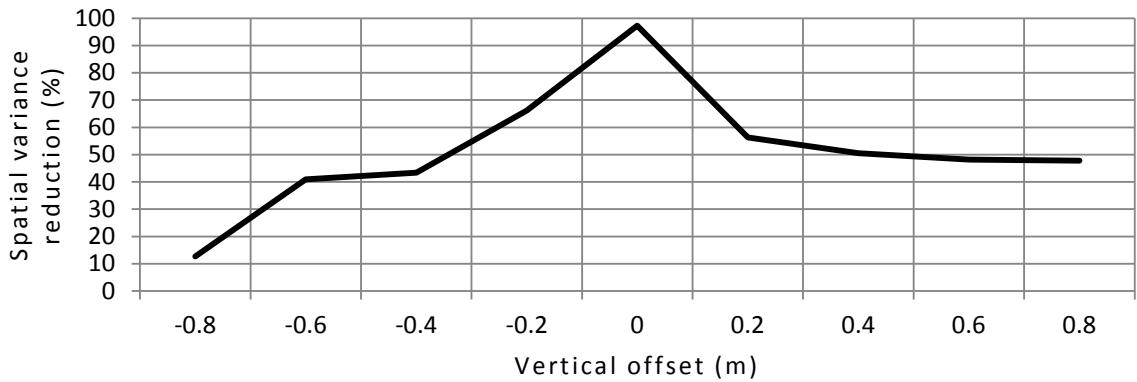


Fig. 5.23 Spatial variance reduction as a function of listening grid vertical offset

The error due to listening grid vertical offset must be considered appropriately, depending on the listener configuration. If listeners are arranged in a seated configuration, it is expected that there will be approximately 40 cm offset in head height due to the varying height of individuals. Based on the results in Fig. 5.23, this means that some listeners will only receive 50% spatial variance reduction while other listeners at the target height will receive over 90%. The goal of correction is to provide an even spatiotemporal response over a wide-area; therefore, this problem must be addressed to give improved correction in the vertical direction while maintaining horizontal performance.

The set of sixteen target points was reconfigured in the form of two concentric cubes. The first cube ($3.0 \text{ m} \times 1.8 \text{ m} \times 1.0 \text{ m}$) contains the first eight target points at each cube corner, where the bottom left and top right corners were chosen as the omnidirectional target points. The inner cube ($1.0 \text{ m} \times 0.6 \text{ m} \times 0.6 \text{ m}$) contains the remaining points, all of which are target points for the full-range system. The overhead view of this setup is displayed in Fig. 5.24.

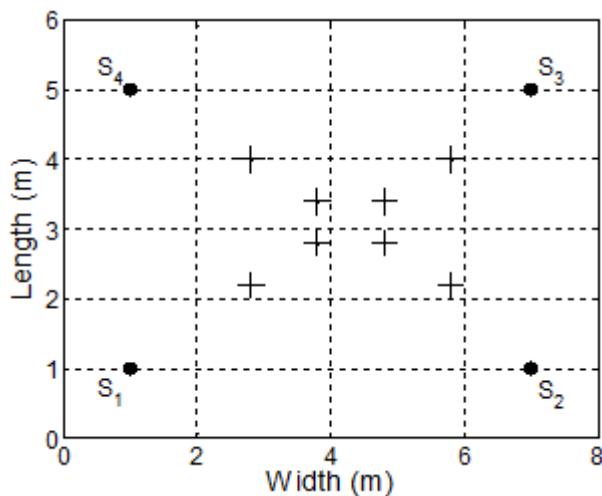


Fig. 5.24 Four-unit CSA system configuration for improved vertical correction (listening grid contains two levels, stacked vertically)

A similar test was performed as in Fig. 5.23, where the 0.0 offset point was set to be the center of the concentric target point cubes which with an offset from -2.0 to 0.8 m. Spatial variance reduction as a function of vertical offset is plotted in Fig. 5.25 for comparison to the non-vertical target point approach shown in Fig. 5.23.

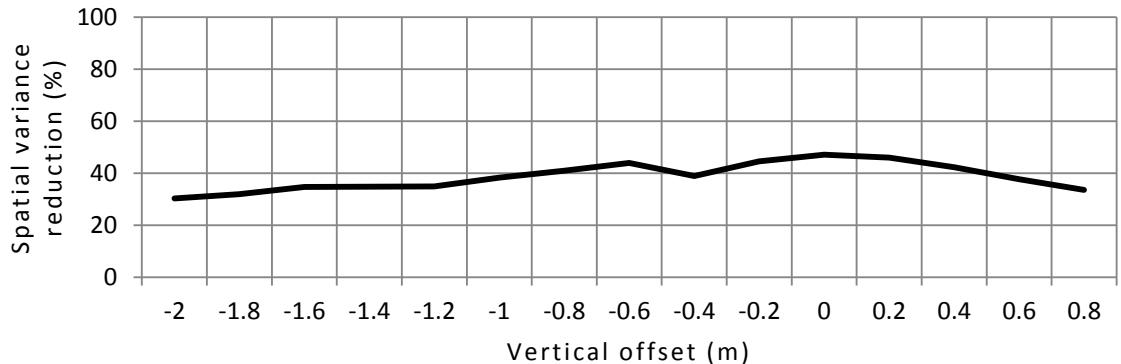


Fig. 5.25 Spatial variance reduction as a function of listening grid vertical offset with the vertically distributed target point grid

The benefits of this CSA configuration are less than with a system targeting only horizontal correction. The mean spatial variance reduction over the vertical range in Fig. 5.25 is approximately 60% while the mean reduction in this case is around 40%. An important point, however, is that even as the vertical range was offset by over two meters the correction benefits did not drop significantly, as was seen in the horizontal-target system. The vertically distributed target point grid provides a consistent response vertically over the listening space which allows for listeners to stand, sit and generally move about without significant distortions to the spatiotemporal response.

It is important to consider the size of listening area each system must correct over. The horizontal correction system addresses a listening area of 6.6 m^2 . The horizontal and vertical system, on the other hand, is responsible for a volume of 5.4 m^3 . This is a considerable difference in spatial correction requirements and therefore the horizontal and vertical correction system should not be expected to perform at par with the horizontal system. If comparable correction performance is desired, more source components and/or target points must be added so that there is a higher density of target points within the listening area to allow for precision control of the acoustical space. This concept is explored in Section 5.5.

Although a wide vertical correction component is necessary in a scenario where listeners are both standing and sitting (and generally moving about the space), it is not entirely necessary if users are seated. This is the assumption for the remainder of this work, unless stated otherwise, as convention dictates that in audiophile listening environments users are generally seated in a precisely laid-out listening area and remain seated for the duration of the audition.

5.4 Performance versus array configuration

The previous section highlights the importance of choosing the appropriate CSA degree and target point placement to achieve the desired correction over a specified area. It is therefore important to understand how physical placement of a CSA can improve the overall correction. This has been touched on in Section 5.3 in regards to dipole unit placement relative to room boundaries. Placing a dipole too close to a reflecting surface results in the reflection (essentially

the inverted direct signal) causing significant destructive interference, sharply reducing the acoustical output level.

This section provides a more comprehensive exploration of CSA unit placement and how a system's passive correction benefits correlate to CSA correction capabilities. A range of practical source layouts are compared with varying source spacing to determine if there is a minimum allowable source separation for proper system operation. Also, source position offset is explored to determine the criticality of maintaining precise CSA unit locations. This is followed by an investigation comparing performance of a hybrid subwoofer CSA to a CSA consisting of single drive-unit subwoofers. This subject is explored with system output efficiency and practicality in mind. The section concludes with a generalized recommendation for overall system configuration of a CSA to ensure the best possible correction benefits along with high system efficiency and practicality.

5.4.1 Array layout performance

CSA simulations indicate a correlation between source layout and correction performance. This section aims to determine a method for estimating the correction benefits of a CSA based on spatial variance reduction due to the passive (uncorrected) layout as compared to a benchmark system. The benchmark for these trials is a single omnidirectional subwoofer in the front left corner of a 5 m x 4 m x 3 m room (20 cm grid point spacing, 10% surface absorption) and inspected over a 36-point walking path (Fig. 5.26).

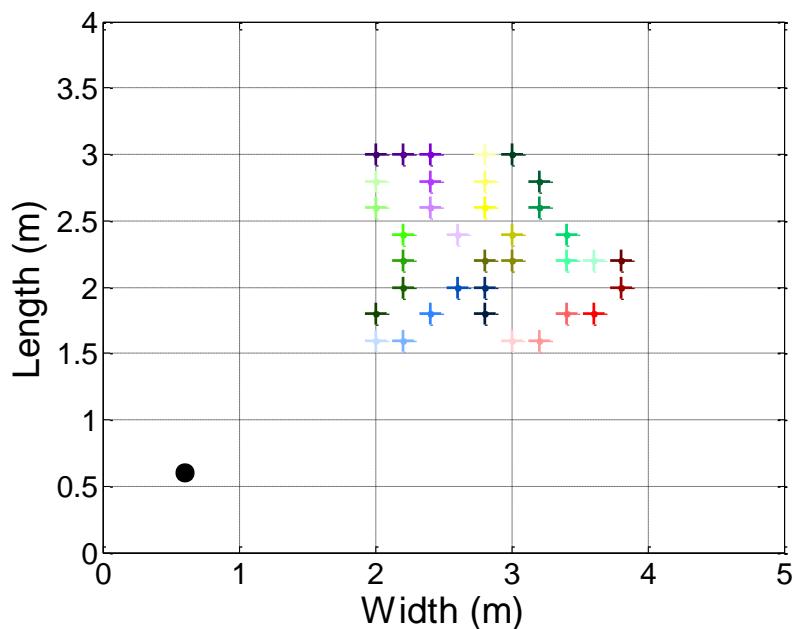


Fig. 5.26 36-point walking path for use with CSA layout testing

Two values are tracked between the uncorrected and corrected systems. First, spatial variance is calculated to determine the uniformity of the frequency response over the walking path. Ideally, this value should be minimized by CSA correction. Second, system output efficiency is deduced by comparing the mean output level (MOL) for the uncorrected and corrected systems. It is expected that efficiency will drop after correction, but reduction should not be excessive for a reasonable system layout.

These two metrics are converted to individual percentage values and then combined to give a net percentage gain in system correction benefits. Each metric is equally weighted, based on the

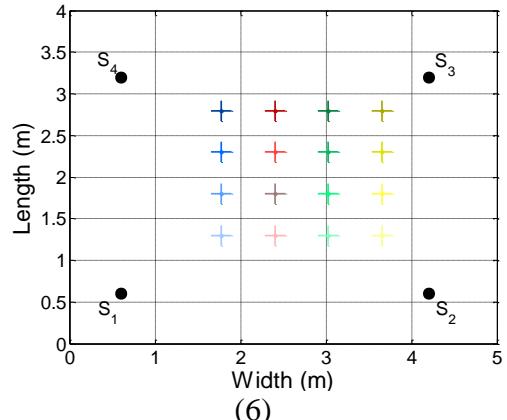
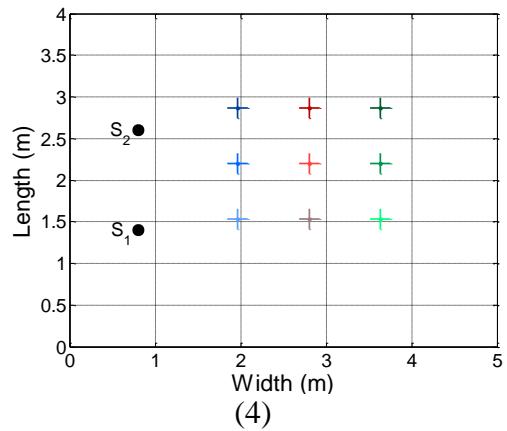
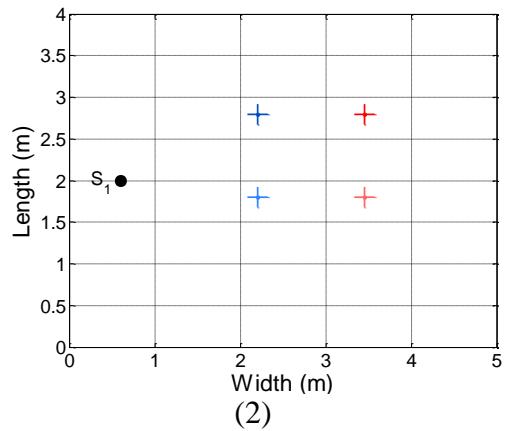
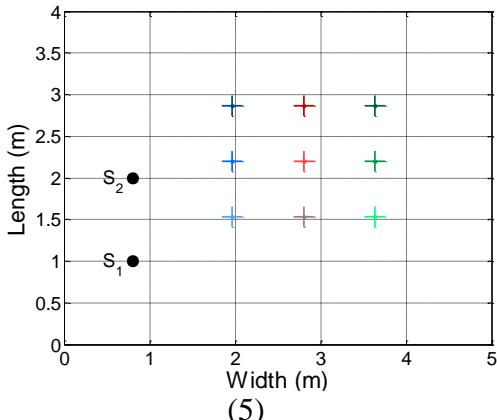
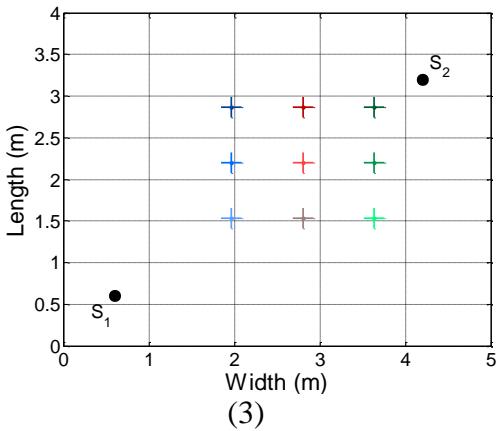
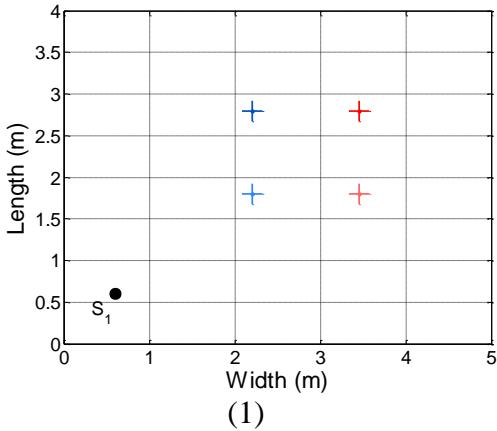
assumption that correction precision is irrelevant without an acceptable MOL post-correction and vice-versa (Eq. 5.7). This method of calculation opens the possibility of negative net gain percentages. These cases indicate that the system actually performs worse (concerning both spatial variance and MOL) after correction.

$$\text{Net Gain \%} = 100 \left(\frac{SV_{UC} - SV_C}{SV_{UC}} - \frac{MOL_{UC} - MOL_C}{MOL_{UC}} \right) \quad (5.7)$$

where:

SV_{UC}	=	uncorrected spatial variance (dB)
SV_C	=	corrected spatial variance (dB)
MOL_{UC}	=	uncorrected mean output level (dB)
MOL_C	=	corrected mean output level (dB)

The tested system configurations are shown in Fig. 5.27 with the collected data in Table 5.3.



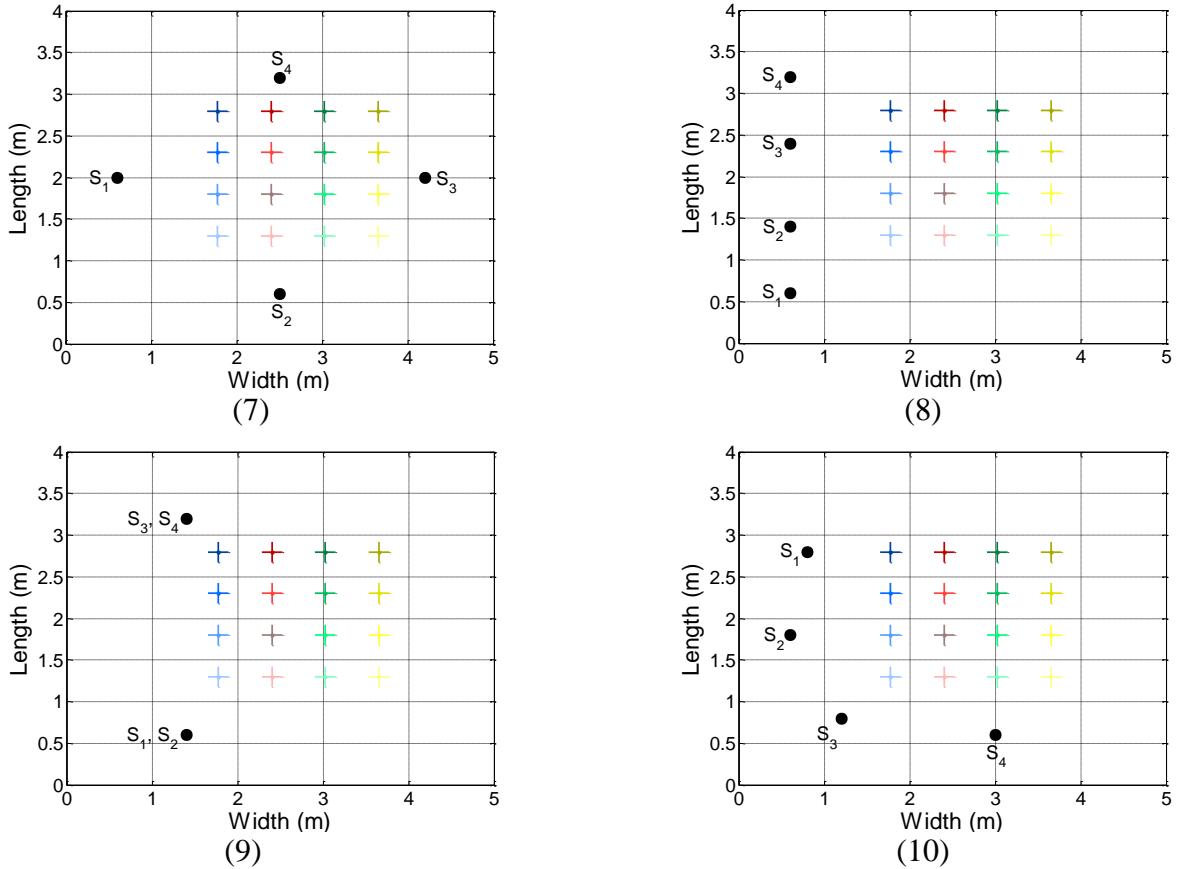


Fig. 5.27 System configurations 1 – 10 for CSA layout testing

#	UC SV (dB)	B to UC SV reduction (%)	C SV (dB)	B to C SV reduction (%)	UC to C SV reduction (%)	Output SPL UC (dB)	Output SPL C (dB)	Output SPL loss UC to C (%)	Net correction gain (%)
1	5.3130	0.0	3.1287	41.1	41.1	72	72	0.0	41.1
2	5.0605	4.8	4.1903	21.1	17.2	75	72	-4.0	17.1
3	3.5508	33.2	1.5477	70.9	56.4	85	70	-17.6	53.3
4	4.8849	8.1	2.3845	55.1	51.2	75	60	-20.0	35.1
5	5.0414	5.1	1.9152	64.0	62.0	80	55	-31.3	32.7
6	3.4057	35.9	0.19684	96.3	94.2	90	83	-7.8	88.5
7	4.4686	15.9	1.2744	76.0	71.5	80	78	-2.5	73.5
8	5.1130	3.8	1.7218	67.6	66.3	80	45	-43.8	23.8
9	5.1403	3.3	0.97854	81.6	81.0	90	75	-16.7	64.9
10	5.0741	4.5	0.80898	84.8	84.1	80	60	-25.0	59.8

Table 5.3 Room configurations for CSA window choice evaluation
(UC = uncorrected, C = corrected, B = benchmark, SV = spatial variance)

Ranking the tested configurations based on net gain (spatial variance reduction minus system efficiency loss) in descending order gives: 6, 7, 9, 10, 3, 1, 4, 5, 8 and 2. The four best performing layouts each encapsulate the target points in some manner. Trials 6 and 7 surround the listening area in the horizontal, XY plane. In trial 9 the sources surround the listening area over the vertical, YZ plane. Trial 10 surrounds the area in a semi-circular manner over the horizontal, XY plane. This aligns with the recommendations proposed in [73] stating that

sources should surround the target listening point. These trials show that adhering to this wisdom does, in fact, provide superior correction.

Another observation is that correction benefits decrease with degrees of freedom. This is more obvious, as it is clear that more target points grant a higher spatial sampling rate, thus tighter system control. There are cases, however, where this principle breaks down, such as in trial 8. The sources are placed in a straight line across the front of the room, resulting in over 60% spatial variance reduction, but also over 40% loss in MOL. This is due to the fact that the source components are all similarly coupled to the room (their responses are not orthogonal), requiring significant processing to achieve the desired response(s).

The best performing configurations have sources at very different locations which have minimal similarities in terms of source to room coupling, providing the desired orthogonal responses for each source component (giving nearly independent degrees of freedom) and resulting in high reductions in spatial variance while maintaining acceptable output levels.

5.4.2 Source spacing

An important subtlety of CSAs is the effect of source spacing. This was alluded to in the previous section where it is seen that systems with sources placed at very different locations perform best. It can be expected that as sources move closer to one another the correlation between each source's room coupling factors will increase, effectively decreasing the independent nature of the system's degrees of freedom.

To highlight the CSA spacing constraints two- and four-unit CSA configurations were tested with source spacing ranging from 0.4 – 5.0 m for the two source system and 0.4 – 3.0 m for the four source system. In each case the first array unit is fixed in the front left room corner to use as a spacing reference for the remaining unit(s). When a source reaches the limit along one room dimension, it continues its spacing path along the perpendicular wall. The spatial variance and efficiency results are presented in Figs. 5.28 & 5.29.

Although the source placement tests are dependent on the target point grid location, a general sense of how CSA unit spacing affects the overall system performance can be deduced. The important value to note is the net gain. This is a sum of the spatial variance reduction and the output efficiency loss percentages to give an overall correction rating.

The two-unit CSA trials (Fig. 5.28) show that although spatial variance can be reduced by upwards of 60% with very small source spacing, the system suffers from over 40% loss in output efficiency. This gives net gains in the range of 20%, which is not significant concerning the goals of a CSA. As spacing exceeds 3.0 m, however, the system stabilizes and there is minimal loss in efficiency, giving a consistent net gain of around 50% for spacing between 3.0 and 4.2 m. This is the range where the sources are maximally independent from one another in terms of source-to-room coupling. When sources are tightly spaced, room-modes are excited in very similar manners, causing difficulties in target response reproduction.

The four-unit trials (Fig. 5.29) exhibit a similar trend to that of the two-unit case. When the units are tightly spaced there is an efficiency loss resulting in a negative net gain after correction. As spacing exceeds 1.8 m, through, the net gain improves to 60% which meets the expectations of CSA correction.

There is a strong dip in correction benefits at 2.6 m source spacing which is dependent on the particular system layout. In this case the dip is a result of a source reaching a room corner (high

modal coupling, thus high spatial variance), where the correction benefits reach a local minimum and then climb once again as source-to-source spacing increases, moving the unit in question away from the room corner.

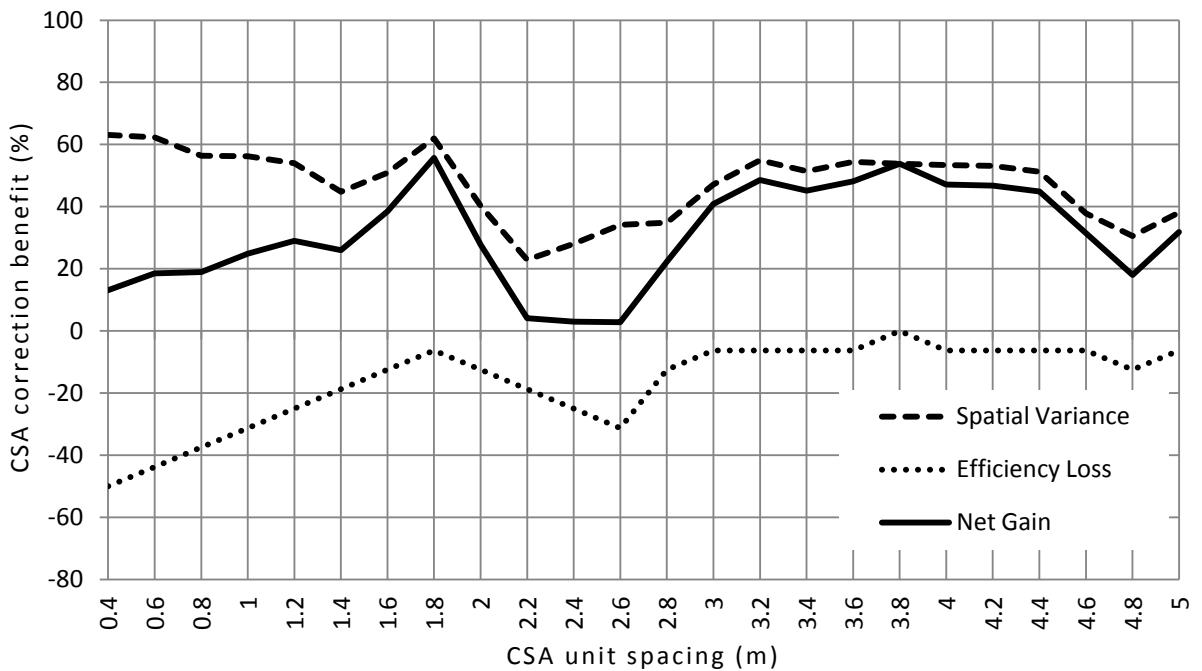


Fig. 5.28 Spatial variance reduction, efficiency loss and net system gain after correction using a two-unit CSA with variable source spacing

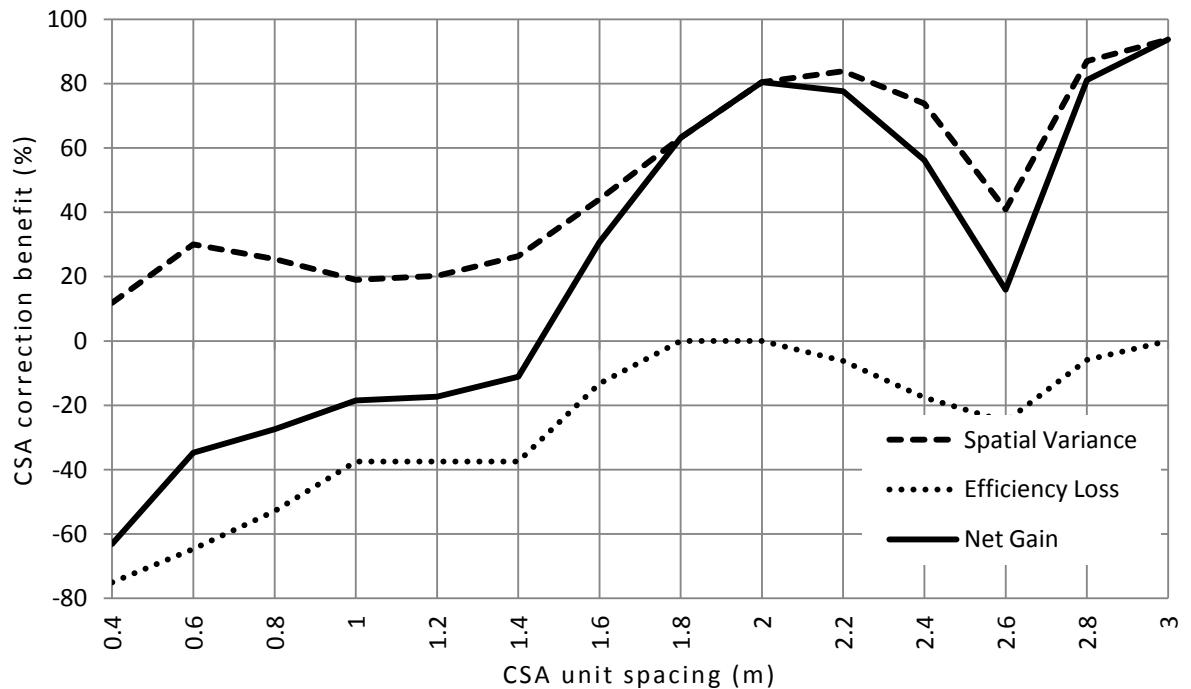


Fig. 5.29 Spatial variance reduction, efficiency loss and net system gain after correction using a four-unit CSA with variable source spacing

5.4.3 Source position offset

While the effects of target point offset have been explored in the form of the virtual walking paths over a listening space, source position offset has yet to be discussed. This section investigates the effects of moving a CSA unit away from its calibration position. Ideally, this spatial offset should not impede the correction benefits over the listening area.

The source position offset tests are conducted in an 8.0 m x 6.0 m x 3.5 m virtual space with grid spacing of 20 cm and surface absorption of 10%. A single-unit CSA is positioned at (2.0 m, 2.0 m, 1.0 m) and a four-point target grid is placed within the space so that correction covers 20 – 120 Hz (Fig. 5.30a). The correction filters are generated and then the single hybrid subwoofer is offset over a range of ± 1.0 m in the x and y dimensions where the resulting spatial variance reduction (from the uncorrected system at the original source placement) is recorded at each point (Fig. 5.30b). The correction filters remain fixed over all positions so that “accidental” subwoofer movement can be analyzed. The two plots are combined in Fig. 5.31 for clarification.

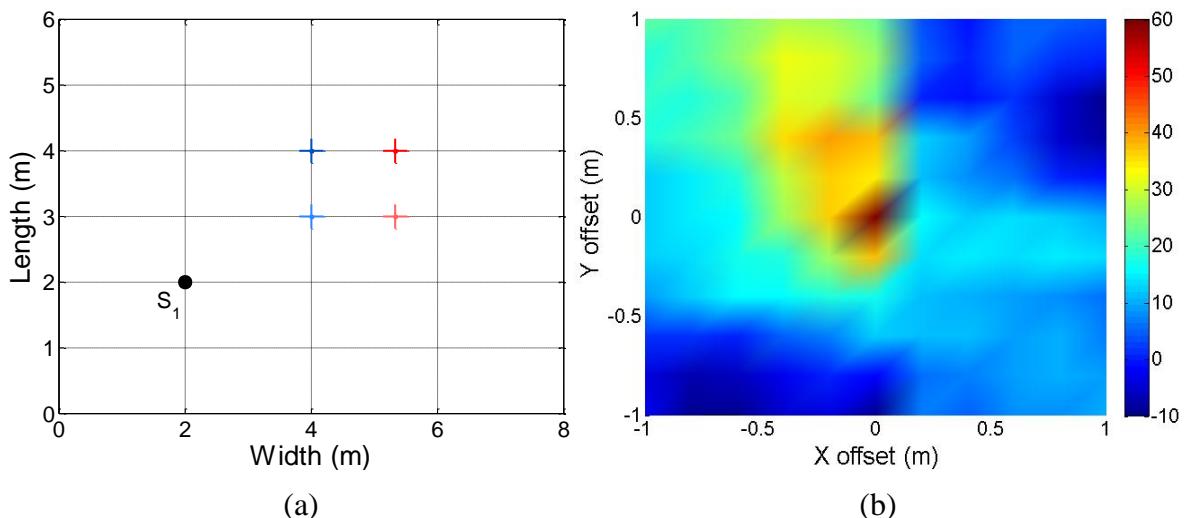


Fig. 5.30 (a) Source position offset test configuration and (b) spatial variance reduction (%) for a single-unit CSA offset by a distance from its intended position (2.0, 2.0, 1.0)

The source position offset testing highlights a number of points concerning CSA layout. Obviously, the system performs best at the intended source position (upwards of 60% spatial variance reduction). As the unit is offset towards the room corner or towards the listening area, though, the correction benefits drop sharply to the point where there is nearly no benefit at all (beyond around 0.4 m offset).

Close proximity to a room corner degrades the system performance since all room-modes have an antinodal point/plane at room corners. As the source moves closer to the corner its coupling to all room-modes increases sharply, invalidating the correction filters, as they were based at a position which was much more loosely coupled to the room-modes. Additionally, as the source approaches the room surfaces, the dipoles lose their efficiency causing changes to the behavior of the system, as discussed in detail earlier in this chapter. Similarly, as the source moves closer to the listening area the source-to-listener coupling changes and causes a loss in correction benefits.

The correction benefits are partially maintained (up to around 40%) when the source is shifted parallel to the room surface and listening area. While coupling factors still change, many

remain constant as the unit is maintaining relative distances across two of the three dimensions. In reality, this is an acceptable characteristic of the correction system as it is likely that if a subwoofer were to be moved it would done so along a wall to make way for a new piece of furniture, for example. In regards to the addition of a new piece of furniture, recalibration is necessary if the size of the piece of furniture introduced is comparable to the shortest wavelength in the subwoofer band (approximately 3 meters or more, such as a large couch). Smaller items aren't likely to have a large effect on the low-frequency response as discussed in [3], therefore recalibration is unnecessary.

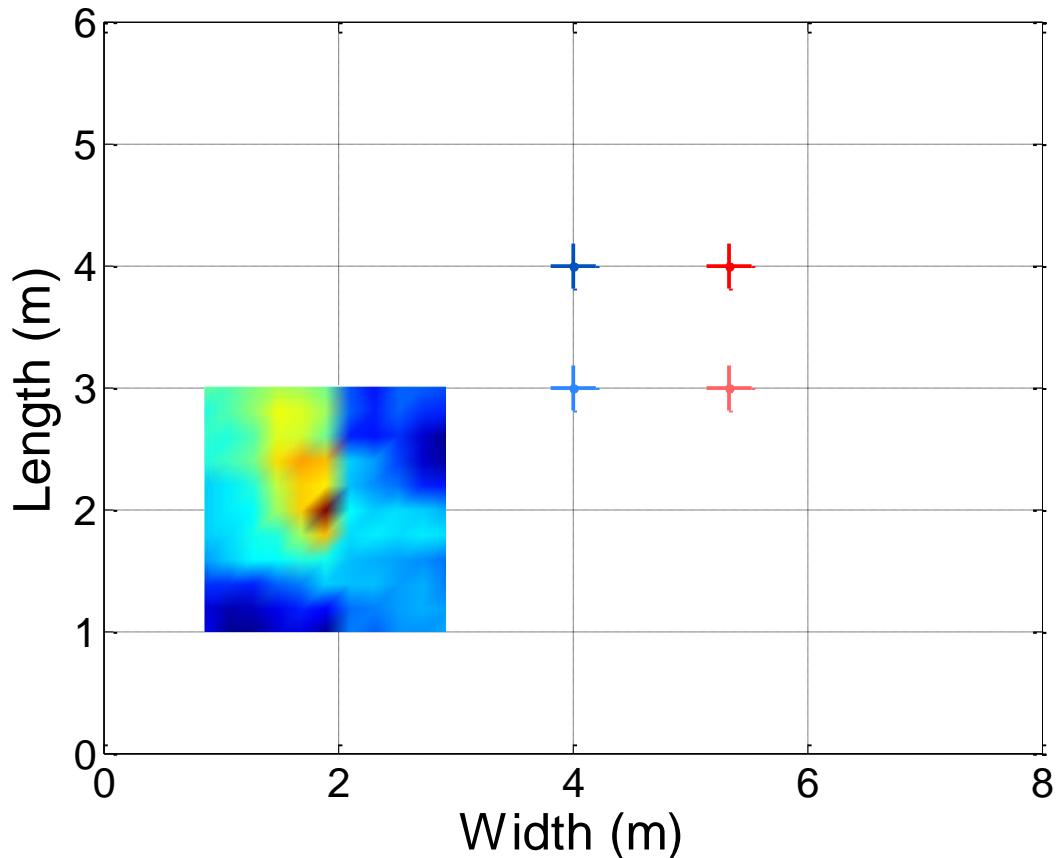


Fig. 5.31 Overlay of spatial variance reduction due to source position offset from (2.0, 2.0, 1.0) and physical system configuration plots

5.4.4 Single versus multiple drive-unit subwoofers

Although the intended implementation of CSA technology involves multiple drive-unit hybrid subwoofers, the correction procedure can be applied to any subwoofer system. A pertinent question therefore is whether systems consisting of multiple single degree of freedom (omnidirectional) subwoofers can match the performance of systems with hybrid subwoofers. In this section hybrid and omnidirectional subwoofer-based systems with equal degrees of freedom are compared based on spatial variance reduction, output efficiency and practicality. This exploration highlights the possibility of CSA correction using existing hardware so that newly introduced system enhancements are only in the form of digital signal processing (DSP).

The first comparison performed involves four degrees of freedom where the hybrid subwoofer system utilizes a single source placed at the room corner (Fig. 5.32a) and the omnidirectional subwoofer system consists of a source at each room corner (Fig. 5.32b). It has been shown in Section 4.1.4 that subwoofers in all room corners provide substantial passive suppression of

room-modes as compared to single unit corner placement. Therefore, the omnidirectional system should naturally give greater correction benefits due to its advantageous configuration.

CSA correction filters were generated as before, except the omnidirectional system operates over a single correction band since the purpose of the dual-band system is to alleviate requirements on dipoles at very low frequencies to boost system efficiency. The virtual space ($5\text{ m} \times 4\text{ m} \times 3\text{ m}$) uses 20 cm grid spacing and 10% absorption. The walking path shown in Fig. 5.26 was used to judge the correction performance. Uncorrected and corrected frequency response plots are shown in Figs. 5.32c-f.

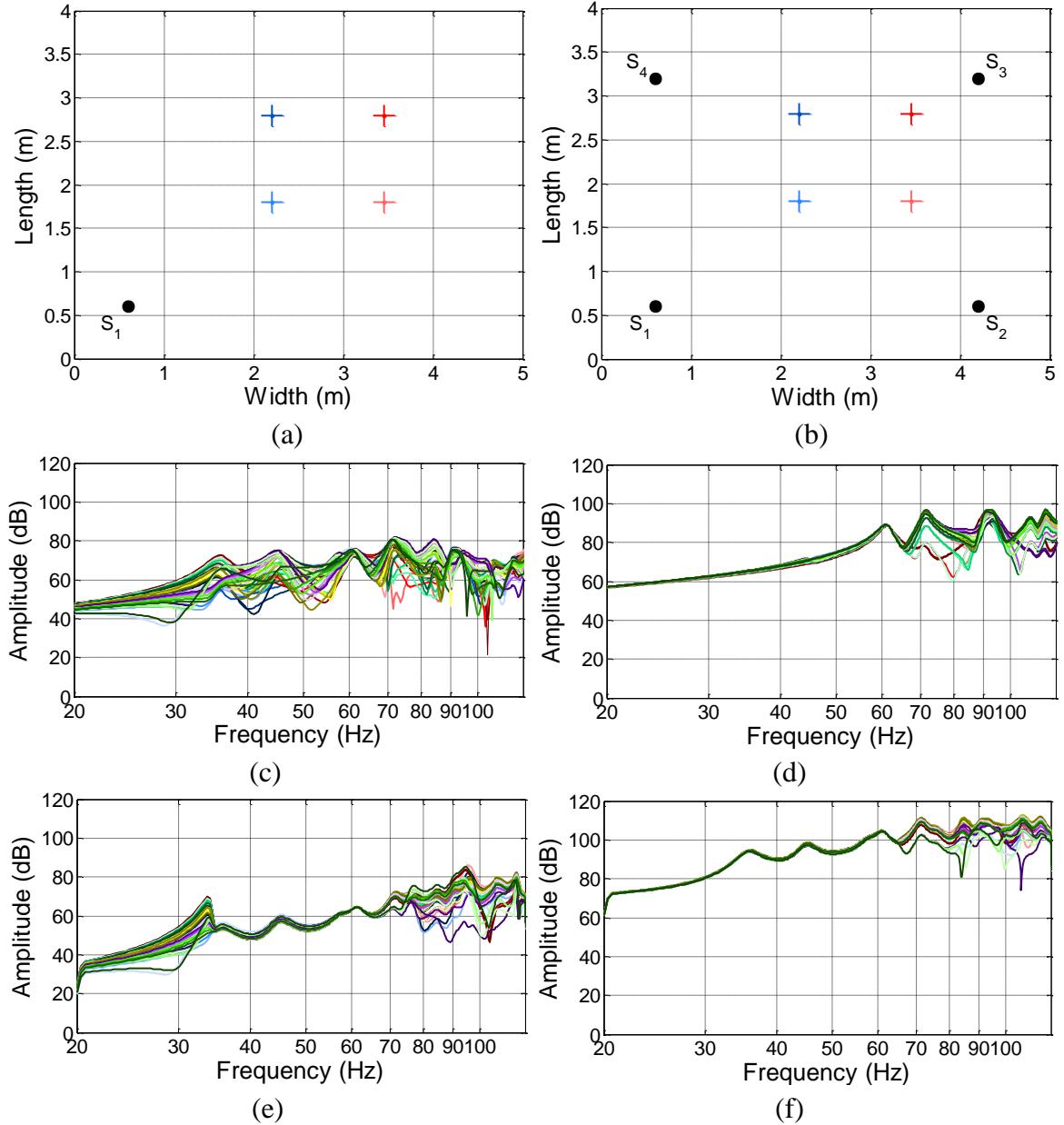


Fig. 5.32 System configuration (top), uncorrected (middle) and corrected (bottom) frequency responses for the single hybrid subwoofer system (left column) and the four-unit omnidirectional subwoofer system (right column)

The hybrid and omnidirectional systems give approximately 30% spatial variance reduction from their respective uncorrected systems. They differ, however, when efficiency is examined. The hybrid system gives no change to system efficiency while the omnidirectional system in

fact gives a 16% boost in efficiency, likely due to a decrease in the passively present destructive interference achieved by phase adjustments in the correction filters. Combining the spatial variance and efficiency benefits, the hybrid and omnidirectional systems give net benefits of 33.1% and 44.5%, respectively. The omnidirectional system's 10% better performance requires three more subwoofers than the hybrid system, which may prove problematic in terms of practicality.

Next, a system comparison is performed between a CSA consisting of the prescribed practical maximum of four units (sixteen degrees of freedom) and a sixteen-unit omnidirectional subwoofer system. The four-unit CSA is placed at room corners so that it benefits from the same passive modal suppression experienced by the omnidirectional system in the previous example (Fig. 5.33a). The omnidirectional system is configured with placement along the perimeter of the room (Fig. 5.33b). The systems were tested in the same configuration as before with the uncorrected and corrected frequency responses presented in Figs. 5.33c-f.

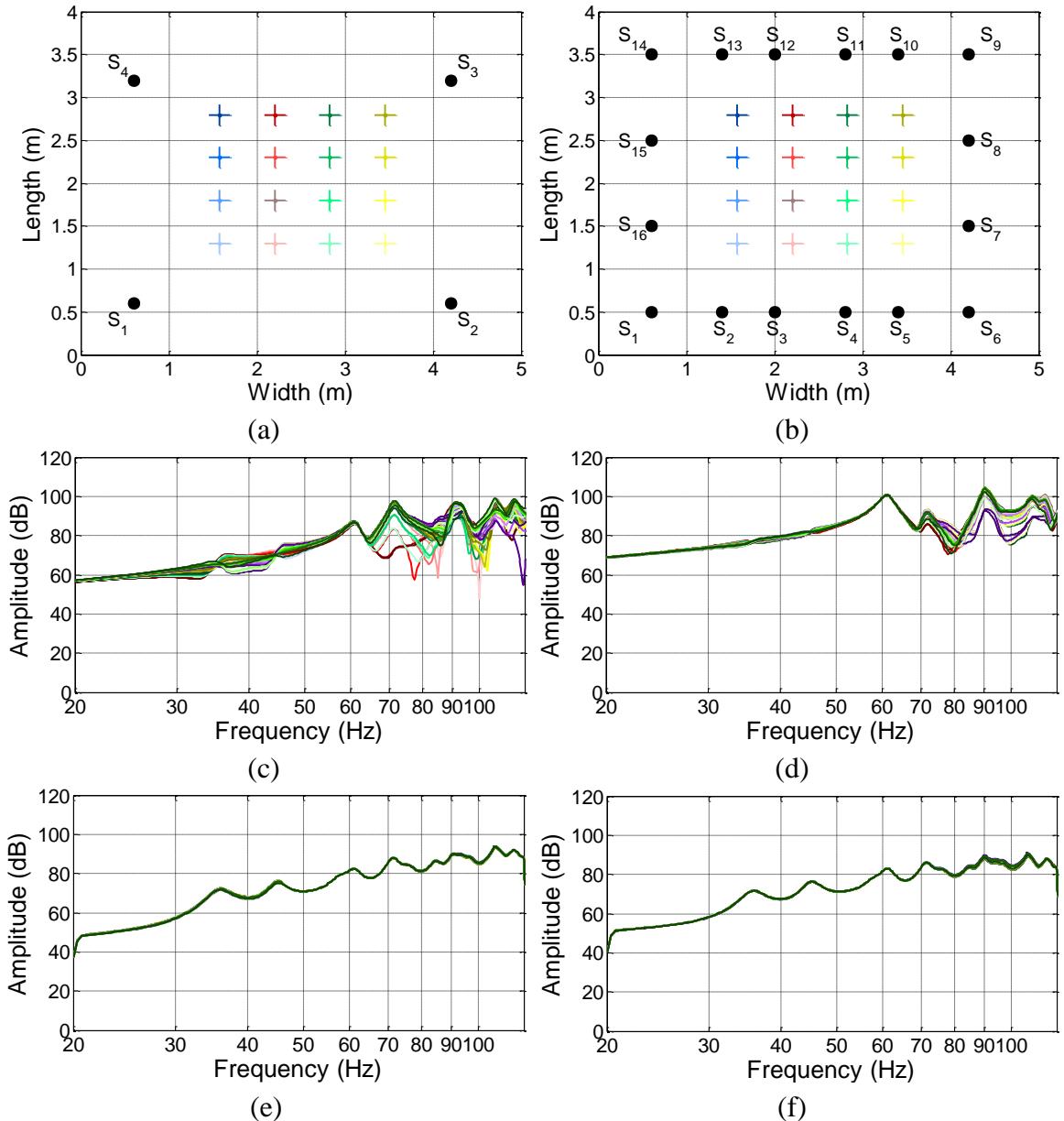


Fig. 5.33 System configuration (top), uncorrected (middle) and corrected (bottom) frequency responses for the four-unit hybrid subwoofer system (left column) and the sixteen-unit omnidirectional subwoofer system (right column)

The comparison between the four-unit hybrid and the sixteen-unit omnidirectional systems highlights the advantages of the multi-component hybrid subwoofers. Disregarding the practical limitations of placing sixteen individual subwoofers in a small-sized listening room, the hybrid system outperforms the omnidirectional system in all areas.

Uncorrected, the hybrid and omnidirectional systems exhibit spatial variances of 2.74 and 1.90 dB, respectively, where the omnidirectional system naturally benefits from extremely low variance due to modal suppression from destructive interference. After CSA correction, spatial variance is largely eliminated over the walking path for both systems, shown by 93.9% and 88.6% reductions for the hybrid and omnidirectional systems, respectively. Both systems experience an efficiency loss of 5.6%. While the two systems' net benefits are in the mid-80 range, the hybrid system rates 5% above the omnidirectional system.

The corrected frequency response plots (Figs. 5.33e-f) are nearly identical both in the shape of the frequency response curves and in amplitude. Although both systems require sixteen independent channels of signal processing and amplification, the hybrid system condenses the physical reproduction components into four relatively small units, as opposed to an impractical sixteen units. This observation dictates that while omnidirectional implementations of CSA correction are equally feasible (and possibly more straightforward) as hybrid unit systems, the large number of individual subwoofers required can prove unrealistic and does not give significant performance benefits over the hybrid-based systems.

5.4.5 Layout recommendations

A generalized set of recommendations for CSA layout can be formulated based on the results presented in the preceding sub-sections. Section 5.4.1 demonstrates the advantages and disadvantages of various physical source layouts. In terms of spatial variance reduction, performance is directly correlated to the degree of the CSA (number of subwoofers) and the passive placement variance characteristics. Higher degree CSAs have advantageous layouts which benefit from higher spatial variance reduction. These systems are subject to high efficiency losses.

Physical layout considerations are explored in Section 5.4.2, where unit spacing is tested with both a two- and four-unit CSA. Both demonstrate that performance is optimized with large source separation. This is attributed to the decorrelation of the respective modal coupling factors for each unit as they move further from one another.

Source position offset from the calibration position causes a reduction in performance. The worst cases of this involve offsets towards a room boundary or the listening area. Offsets parallel to a boundary/listening area cause much less severe impairments to the correction. Generally, offsets above 40 cm in any direction cause significant degradation of the CSA benefits, but lesser offsets do not have as drastic an effect.

Lastly, it is possible to utilize omnidirectional subwoofers as opposed to the proposed hybrid, multi-component subwoofers. These systems are shown in Section 5.4.4 to function equally well within the CSA correction algorithm. To achieve the necessary degrees of freedom these omnidirectional systems may require an impractical number of subwoofers, thus leading to the recommendation of hybrid subwoofer CSAs.

In summary, CSAs show optimal operation when placed in positively behaving passive configurations with maximal unit spacing. Omnidirectional subwoofers can be used, although hybrid subwoofers are recommended to keep hardware requirements in the practical range.

While it is not recommended to move a source after system calibration, small position offsets of up to 40 cm are tolerable, although system re-calibration is preferred to ensure the best possible performance over the listening area.

5.5 Additional target points and pseudo-inverse filtering

The preceding sections of this chapter require CSA source component and target measurement point quantity equality in order to directly calculate the inverse matrix (Eq. 5.1). It is possible, however, to provide CSA correction to systems containing an unequal number of source components and target points. The Moore-Penrose pseudo-inverse [75] provides a technique for determining the solution to a system of linear equations that would otherwise be impossible due to the non-square matrix in question. The pseudo-inverse operates by finding the best fit solution for the matrix inverse, commonly with a least-mean squares (LMS) algorithm. This function is available in MATLAB as the *pinv* command.

As alluded to earlier in this chapter, the addition of extra target points gives an improved spatial sampling of the listening area and therefore may provide a more robust solution for CSA correction while avoiding a target point resting on a multi-nodal point having strong influence on the correction filters. Additionally, if a simpler correction procedure is required, system calibration can be expedited by utilizing fewer measurements.

5.5.1 Variable listening area, fixed target point density

The pseudo-inverse technique was first applied to a scenario with a fixed number of source components (in a static configuration) using a variable listening area size, but with said area covered uniformly with target points. A virtual 5 m x 4 m x 3 m room was setup in the FDTD toolbox with 20 cm grid point spacing and 10% surface absorption. All target points were placed at each grid point in the listening area, ensuring a fixed target point density. More target points allow for larger listening areas, hence the variable listening area descriptor of this section.

Four system configurations were tested: one subwoofer in the front left room corner, two subwoofers in opposite room corners, three subwoofers in room corners and four subwoofers in room corners. All subwoofers were of the hybrid, four degree of freedom variety. Each configuration had CSA correction applied using a varying number of target points. The number of points used for each trial was: 4, 9, 16, 25, 36, 49, 64, 81 and 100. Spatial variance reduction between the uncorrected and corrected systems was recorded in each permutation over a predefined walking path over the listening area (Fig. 5.34).

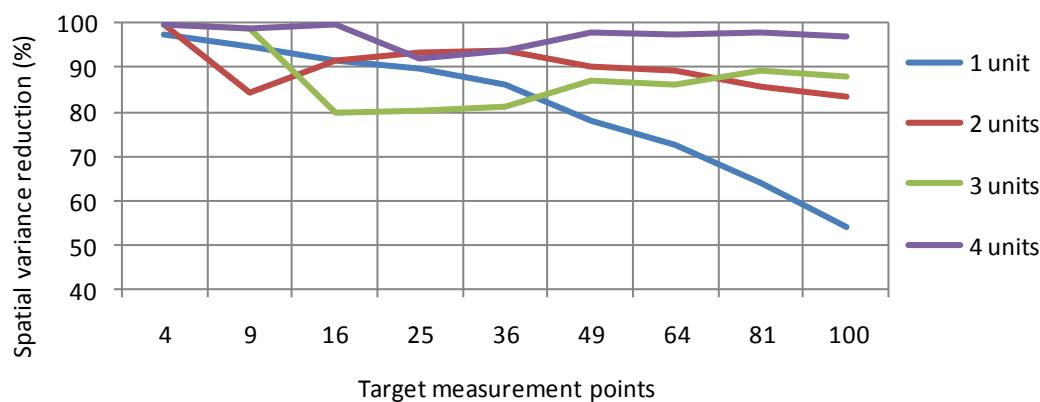


Fig. 5.34 Spatial variance reduction as a function of the quantity of uniformly-spaced target measurement points (variable listening area size)

This set of trials enforces the usefulness of a high number of degrees of freedom. The two-, three- and four-unit CSAs perform consistently, regardless of listening area size. The pseudo-inverse processing gives adequate correction for all points, despite the fact that there are an insufficient number of source components to directly address each target point.

Problems arise with the single-unit CSA. This system ideally operates with four-target points and indicates over 95% spatial variance reduction in Fig. 5.34. As the quantity of target points (and the corresponding listening area size) increases, the effectiveness of the CSA declines until at one-hundred target points spaced over a 2 x 2 m area only 55% spatial variance reduction is achieved. This is a considerable loss from the ideal four-point case and indicates that the frequency responses over larger targeted areas differs sufficiently and a robust pseudo-inverse solution cannot be found using only four source components.

5.5.2 Fixed listening area, variable target point density

It is clear that CSAs of two or more hybrid subwoofers are capable of correction over a wide listening area with uniform target point spacing; however it is also important to understand the relationship between target point density and correction capabilities over a fixed listening area. This section tests this concept by defining a fixed 2 m x 2 m listening area in the same virtual space utilized in Section 5.5.1.

The quantity of target points is governed by how many fit within the area based on the defined spacing. The maximum spacing utilized is 2 m, which corresponds to four target points. Minimum spacing is 20 cm, which requires one-hundred target points. Spatial variance reduction between the uncorrected and corrected systems is displayed in Fig. 5.35.

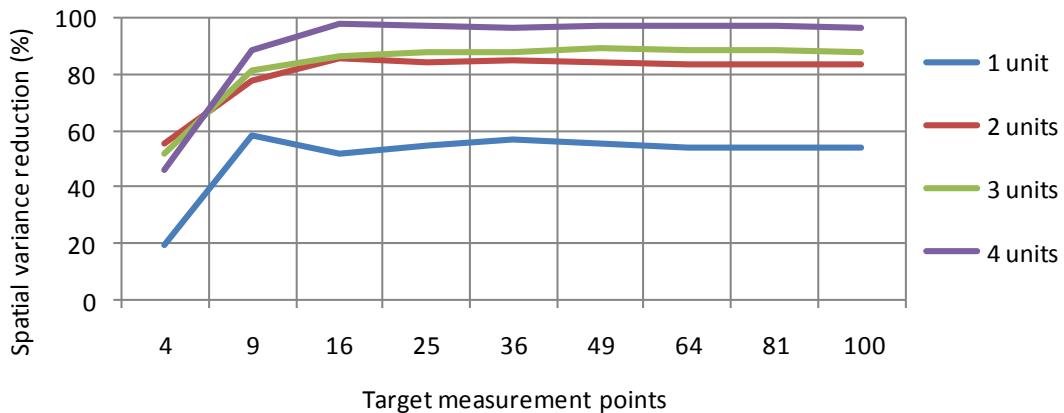


Fig. 5.35 Spatial variance reduction as a function of the quantity of variably-spaced target measurement points (fixed listening area size)

All CSA configurations show similar trends in this set of trials. When four target points are utilized (2 m point-to-point spacing) the correction benefits suffer. As target spacing increases, though, the benefits stabilize where at sixteen target points or greater, there is little change in spatial variance reduction. The point-to-point spacing with sixteen points is 60 cm, corresponding to a quarter-wavelength at 140 Hz. This is close to the upper limit of the defined subwoofer band (120 Hz) and conforms to the requirement of quarter-wavelength target point spacing for accurate correction.

Above sixteen points (less than 60 cm spacing) the system oversamples the listening area, where additional points are not necessary. The original target point spacing criteria presented in Section 5.2 is therefore reasonable for CSA correction and is valid for CSAs of any degree.

5.5.3 Pseudo-inverse filtering conclusions

Pseudo-inverse filtering is beneficial within the CSA correction algorithm. It is likely that practical implementations of CSA technology will have a limited (and not ideally high) number of source components. In the direct one-to-one source component to target point approach, this leads to severe limitations on correction, where a CSA may perform no better than other existing room-mode correction methods.

Eliminating the one-to-one restriction using pseudo-inverse processing allows users to take additional measurements to conform to the quarter-wavelength spacing criteria for accurate correction. The simulations conducted in this section illuminate the advantages of this approach while also discouraging spatial oversampling.

The first set of trials show that as the listening area expands correction benefits decrease for systems with few degrees of freedom, even though target points were uniformly spaced at 20 cm (which should allow for accurate correction beyond 120 Hz). Systems with eight or more degrees of freedom, however, deliver consistent correction over a randomized walking path.

A fixed listening area with variable target point spacing (and the quantity of target points) highlights the need for quarter-wavelength spacing. As point spacing increases beyond the quarter-wavelength limit (60 cm, in this case), spatial variance reduction lowers by approximately 40% in each case, regardless of the available degrees of freedom. At quarter wavelength spacing or better the correction benefits are consistent leading to the conclusion that target points must be tightly spaced to handle the entire subwoofer frequency range, even if this means adding more points than there are degrees of freedom. This ensures the space is adequately sampled and the CSA will perform optimally.

Pseudo-inverse processing allows for system flexibility where any number of measurements can be used in order to provide the most accurate correction. This concept strengthens the argument for an upper array degree of four, as trials in this section indicate that adequate correction over the entire subwoofer band is feasible with only four hybrid units using additional target measurement points and pseudo-inverse processing.

5.6 CSA prototype development

The theoretical and virtual development of CSAs presented in the preceding sections of this chapter provides a thorough blueprint for a CSA prototype. Although various inconsistencies between simulations and real-world implementations are expected due to nonlinearities present in practical environments, the implementation suggestions should allow for a prototype capable of spatial variance reduction over a wide listening area. This section details this development, first using an omnidirectional subwoofer CSA and later a CSA consisting of a single hybrid subwoofer.

5.6.1 Omnidirectional subwoofer CSA experimentation

Prior to development of a hybrid subwoofer-based CSA, an omnidirectional subwoofer CSA was implemented. This particular system consists of two different makes/models of subwoofers: two Bowers & Wilkins ASW 750 and two KEF PSW 1000.2. Both subwoofer types are active, therefore no external amplification is necessary. Independent drive signals are sent to each subwoofer via a MOTU 2408 [76] external sound card which has eight analog audio outputs. All calibration measurements are taken using a customized GUI developed within MATLAB (Fig. 5.36) while correction evaluation measurements are taken with a CLIO

SC-02 [77] box and the accompanying software along with an Audiomatica MIC-01 measurement microphone. The system was configured as shown in Fig. 5.37. The subwoofer and target measurement locations for the initial testing are indicated by the small green boxes and the large central red box, respectively. The black boxes in the figure represent non-rectangular aspects of the test space.

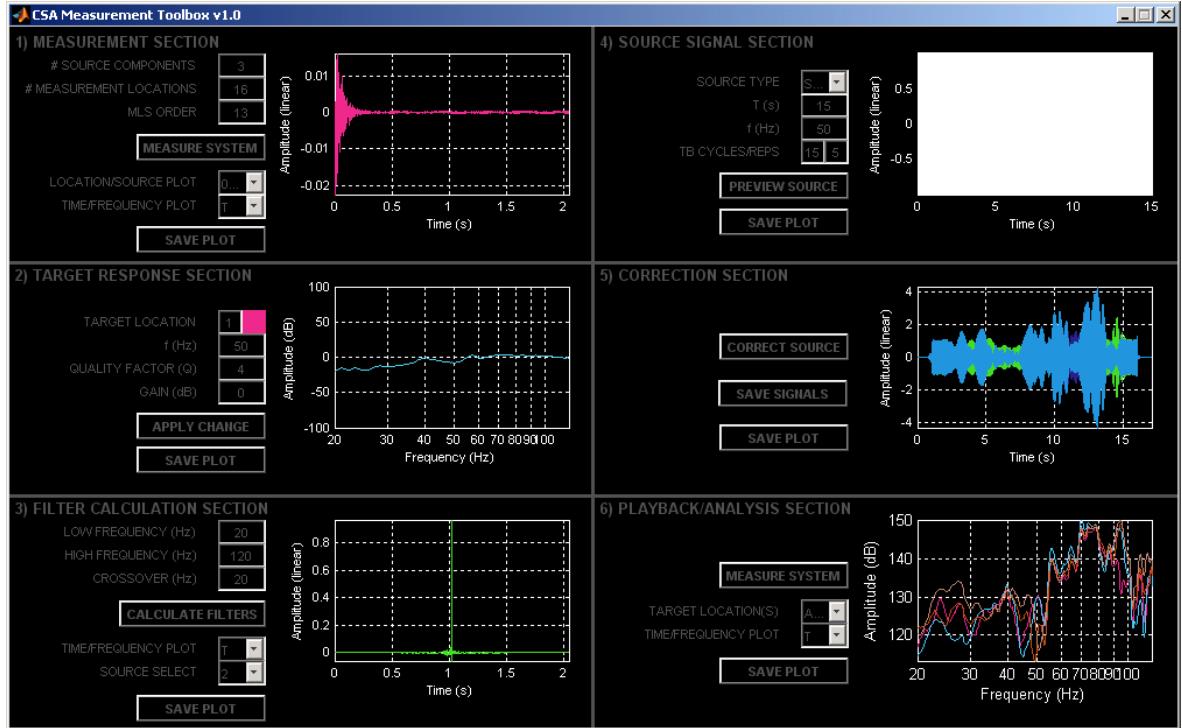


Fig. 5.36 CSA measurement toolbox GUI screenshot

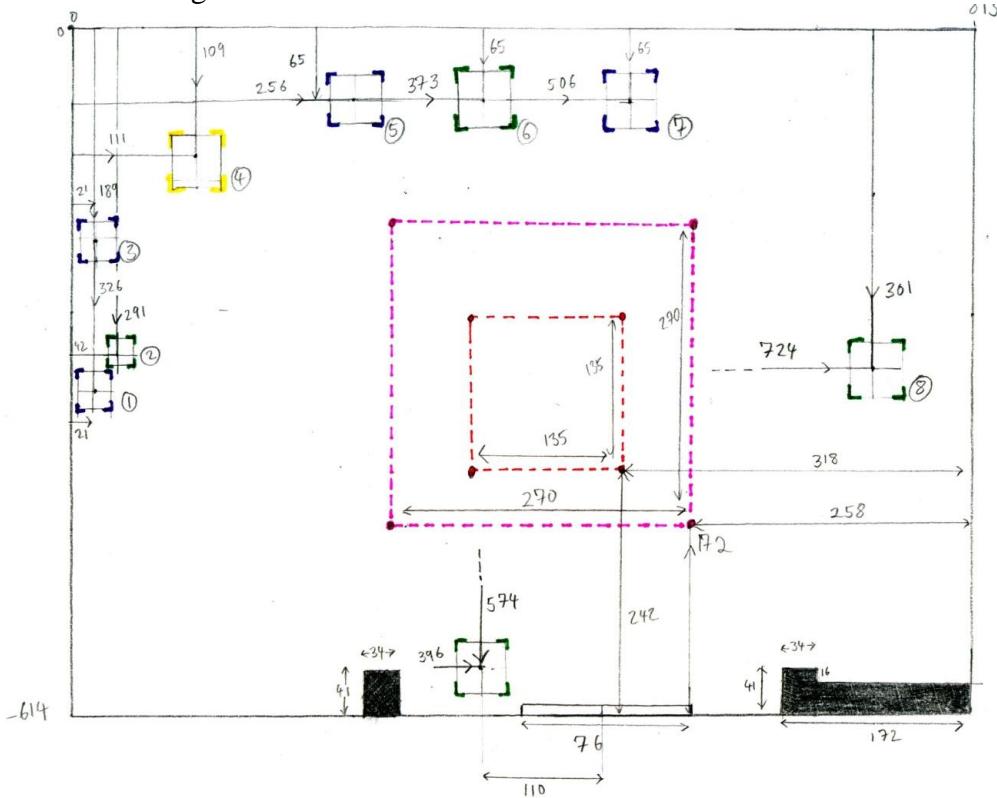


Fig. 5.37 CSA experimental configurations (all values in cm, room height = 2.74 m, target point height = 1.60 m, subwoofers placed on the floor)

As with the simulations, the experimental CSAs require calibration measurements. The system was first tested over a 1.35 m x 1.35 m listening area consisting of 100 target points spaced at 15 cm intervals throughout. Although this many target points is unnecessary (as demonstrated in Section 5.5) they were included to ensure the target area was properly sampled to provide the best possible correction. After the target point impulse responses were calculated (using MLS exactly as with the simulations) the system average response was determined (Fig. 5.38) which will serve as the target response.

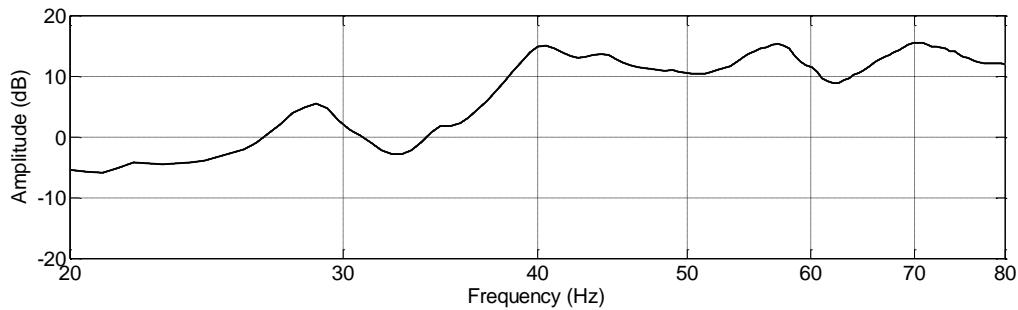


Fig. 5.38 Target frequency response for the measured omnidirectional CSA

The measurement and target matrices were fed through the CSA algorithm to generate the correction filters which were then applied to a five second logarithmic swept sinusoidal signal for system evaluation. To best judge the correction benefits the system was simulated using the FDTD simulation toolbox. The uncorrected and corrected systems were each evaluated over an identical walking path, as shown in the configuration plot (Fig. 5.39). The measurement and simulation results are displayed in Figs. 5.40 and 5.41, respectively.

This system configuration naturally gives low spatial variance due to the wall midpoint placement of the subwoofers (as discussed in Chapter 4.1.4). In both the measured and simulated cases, CSA correction provides additional spatial variance reduction. The measured system gives 59.5% reduction while the simulated system gives 52.5% which are in close agreement to one another, providing further validation for the FDTD toolbox.

The measurements and simulation differ, however, in their uncorrected frequency responses. These differences can be largely attributed to differences between the two types of subwoofers used for the measurements and the modeled point-source subwoofer utilized in the simulations. The different frequency responses of the subwoofers in the measured system combine to give the overall response differently than with a system consisting of identical subwoofers. This disparity is not modeled in the FDTD toolbox, so differences in simulated and measured responses should not be unexpected.

Additionally, the FDTD toolbox models all surfaces as perfectly rigid with frequency-independent absorption. The measured environment, on the other hand, contains a wide-range of surfaces. Three of the four walls and the floor are highly-reflective rigid materials (with parts of the front, left and right walls covered with heavy curtains). The upper portion of the front wall is a large glass window which is likely to exhibit reactive properties. The ceiling consists of a grid of acoustical tiles and ventilation ducts. Most of these elements are not included in the acoustical model, and therefore contribute to the inaccuracies between the measurements and simulations. The most noticeable divergence is a large 65 Hz attenuation in all measurements (Figs. 5.40, 5.43 & 5.46) which can be presumed to be caused by the non-rigid glass surface across the upper half of the front wall, as the dip is present in all measurements, regardless of the subwoofer configuration.

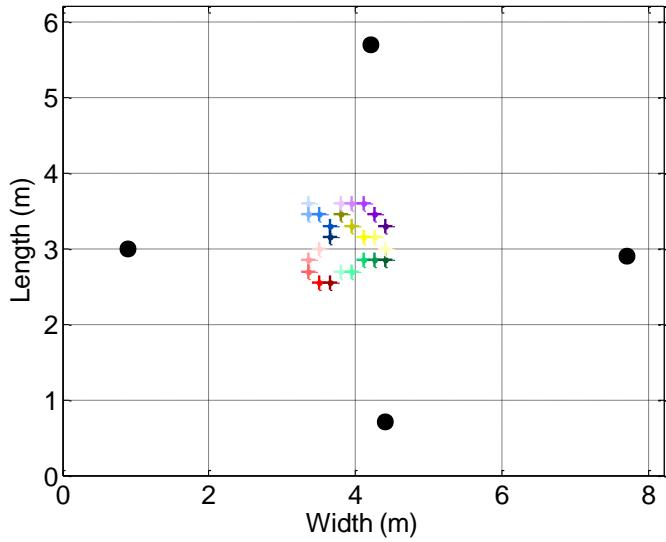


Fig. 5.39 Simulated and measured system configuration (including the walking path)

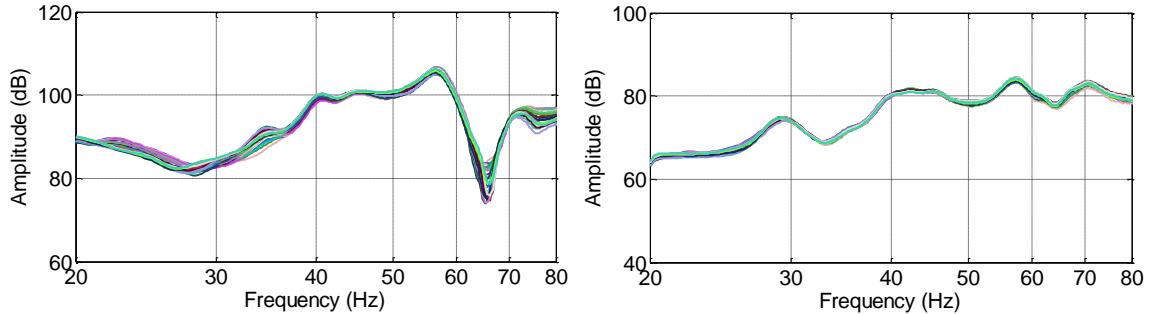


Fig. 5.40 Uncorrected (left) and corrected (right) omni-subwoofer CSA measurements

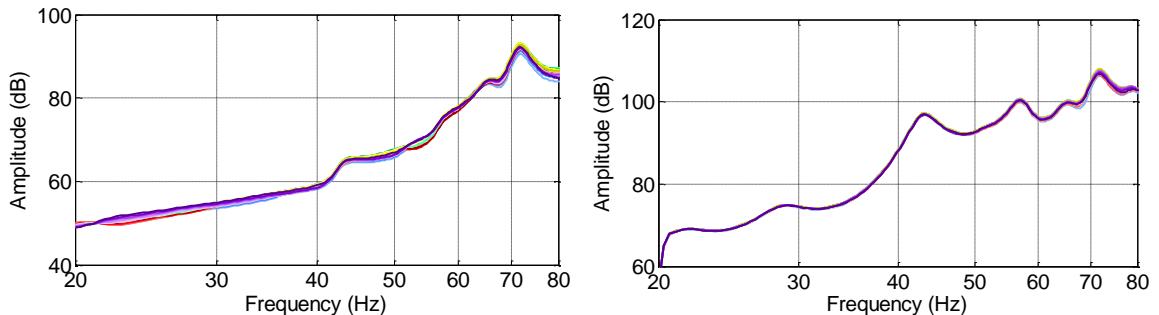


Fig. 5.41 Uncorrected (left) and corrected (right) omni-subwoofer CSA simulations

The measured and simulated corrected frequency responses do, on the other hand, exhibit similarities. The target frequency response is much more reliant on a space's acoustical characteristics rather than the subwoofers' attributes since the target response is an averaging of responses across the listening area. The measured target response (Fig. 5.38) nearly matches the measured responses in Fig. 5.40. This target response is extremely similar to the post-correction simulated responses in Fig. 5.41. Although the measured room and the modeled room are not identical due to approximations within the simulation structure, the basic characteristics of the responses indicate that the two methods are in close agreement.

The test was repeated with a listening grid consisting of 100 target points but with 30 cm spacing, corresponding to a listening area of 2.7 m x 2.7 m (the central pink box in Fig. 5.37).

The system configuration along with the measured and simulations are presented in Figs. 5.42 – 5.44, respectively.

As with the tighter-spaced listening grid, the wider listening area benefits from CSA correction in a similar manner, experiencing measured and simulated spatial variance reduction of 51.1% and 60.3%, respectively. The measured and simulated responses differ due to the reasons already specified. The wider target point spacing demonstrates the reduction of correction benefits at high frequencies due to a lower spatial sampling rate. Nevertheless, CSA correction provides accurate control by achieving the target response over the listening area.

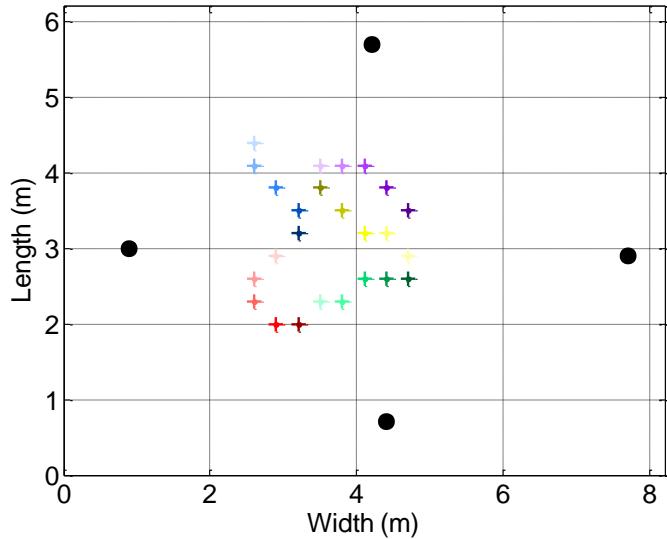


Fig. 5.42 Simulated and measured system configuration (including the walking path)

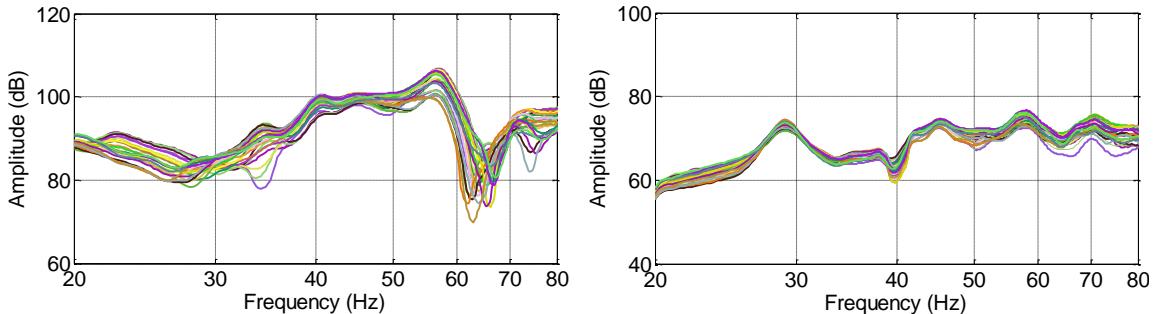


Fig. 5.43 Uncorrected (left) and corrected (right) omni-subwoofer CSA measurements

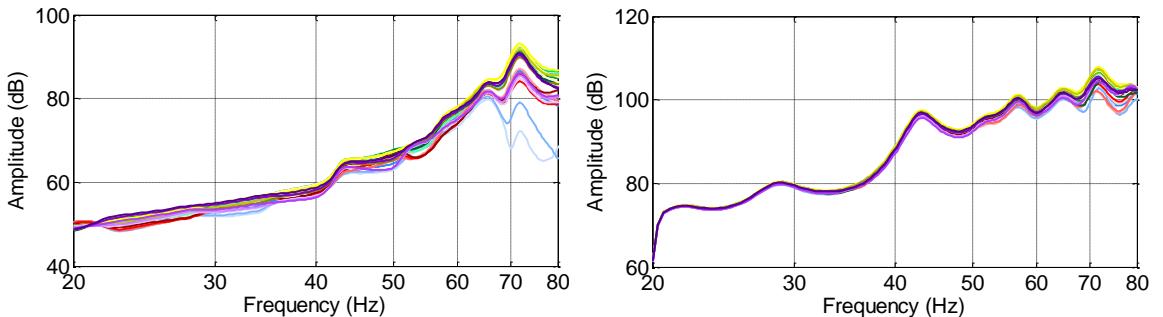


Fig. 5.44 Uncorrected (left) and corrected (right) omni-subwoofer CSA simulations

Lastly, the CSA was arranged in a more random configuration where the subwoofers were no longer in “ideal” positions (Fig. 5.45). The listening grid utilized is the same as with the first test (15 cm spacing) and the results are shown in Figs. 5.46 and 5.47.

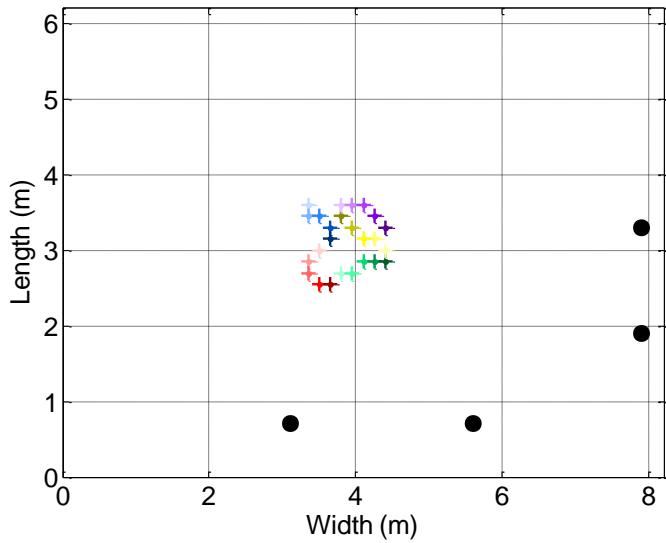


Fig. 5.45 Simulated and measured system configuration (including the walking path)

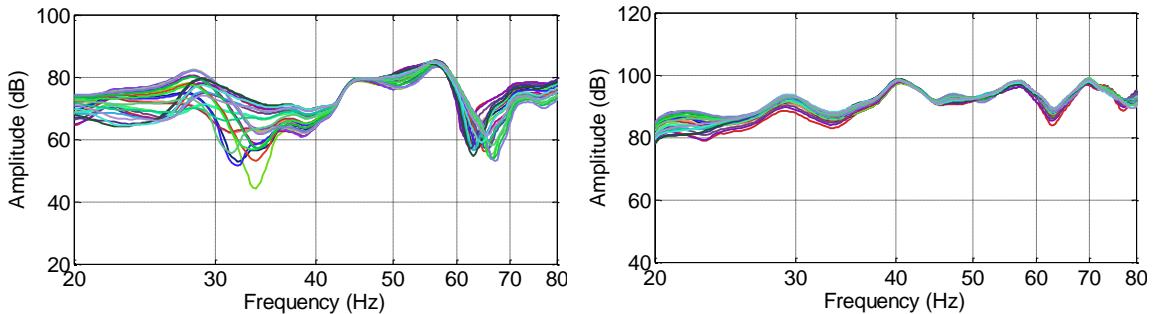


Fig. 5.46 Uncorrected (left) and corrected (right) omni-subwoofer CSA measurements

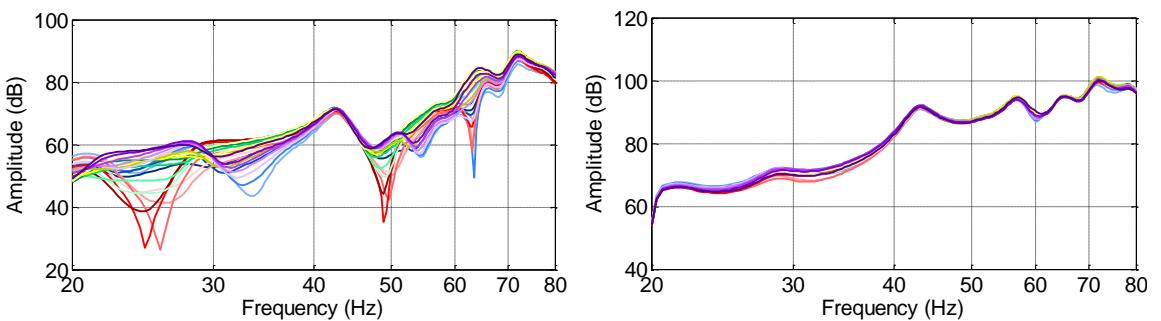


Fig. 5.47 Uncorrected (left) and corrected (right) omni-subwoofer CSA simulations

This test configuration gives 68.3% and 80.8% measured and simulated spatial variance reduction, respectively. While the simulation performs better than the measured system, both CSAs still provide considerable improvements to the response across the listening area. Additionally, as the subwoofers aren’t placed at positions which cause high destructive interference (as with wall midpoint placement), system efficiency is improved, as seen in the similar output levels between the uncorrected and corrected systems.

5.6.2 Prototype hybrid subwoofer development

To demonstrate the practicality of the proposed CSA implementation, it was necessary to develop a hybrid subwoofer prototype. This task was carried out in two stages. The first stage was part of two final-year undergraduate projects where students constructed a subwoofer with an independently-controllable drive-unit on each enclosure face (Fig. 5.48).



Fig. 5.48 Initial construction of the hybrid subwoofer prototype

This initial construction had two primary design issues. First, the enclosure dimensions (55 cm cubed) were impractical. The large enclosure faces allowed for high structural resonance which was evident during performance testing, especially when opposing drive-units were moving in contrary directions. Secondly, the enclosure faces were poorly sealed, causing subjectively disturbing whistling due to air transmission.

Despite these drawbacks, the students achieved approximate low-frequency polar pattern control (omnidirectional and rotated dipolar and cardioid patterns) in outdoor tests. It was concluded, however, that this construction would not be beneficial for CSA applications due to the highlighted drawbacks, necessitating a redesign (the second stage of development) as shown in Fig. 5.49.

Three areas were addressed for the redesigned subwoofer. First, the enclosure dimensions were reduced to a 44 cm cube; this was the smallest size possible due to internal spacing constraints of the six drive-units. Secondly, internal bracing was installed between opposing enclosure faces to provide increased structural rigidity when drive-units move in opposition. Lastly, silicone-based sealant was applied to the wall junctions to provide a seal so that the undesirable whistling was eliminated. The redesigned hybrid subwoofer is used for the CSA experimentation in the following section.



Fig. 5.49 Redesigned hybrid subwoofer used for practical CSA experimentation

5.6.3 Hybrid subwoofer CSA experimentation

The redesigned hybrid subwoofer was configured to operate within a CSA correction system. The subwoofer configuration does not precisely match that used in simulations due to limitations of the measurement software in MATLAB which allows for only one pair of outputs on the MOTU sound card to be active at one time. While it is possible to access each available sound input/output device connected to a computer within MATLAB, the built-in functions only allow simultaneous playback/recording on a single stereo pair of inputs and outputs at one time. MATLAB sees the MOTU 8-channel sound card as a series of four stereo sound cards. This prohibits utilizing all drive-units for the omnidirectional source component without rewiring the system. Since the ideal hybrid subwoofer configuration requires driving all units synchronously for the omnidirectional pattern, but independently for the dipole patterns, a compromise is required for the current prototype (although a workaround in MATLAB is likely possible).

The prototype system's four degrees of freedom were arranged as follows: two independent “omnidirectional” components (the top and bottom drive-units) and two dipolar pairs (one for each horizontal axis). The hybrid subwoofer CSA was tested using the 15 cm target point spacing (Fig. 5.50) with the measurement and simulations results shown in Figs. 5.51 and 5.52, respectively. Note that the hybrid subwoofers in the experimental and simulated systems are not identical. The simulation utilizes the source components as dictated earlier in this section (one omnidirectional and three dipolar components); therefore the uncorrected frequency responses may not be in close agreement.

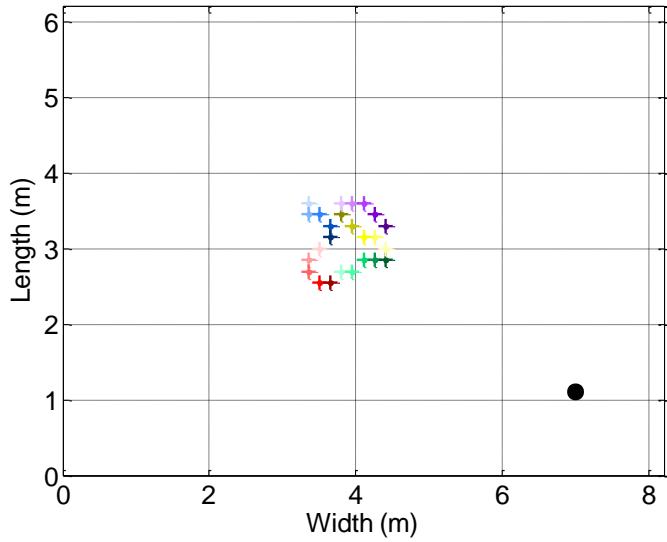


Fig. 5.50 Simulated and measured system configuration (including the walking path)

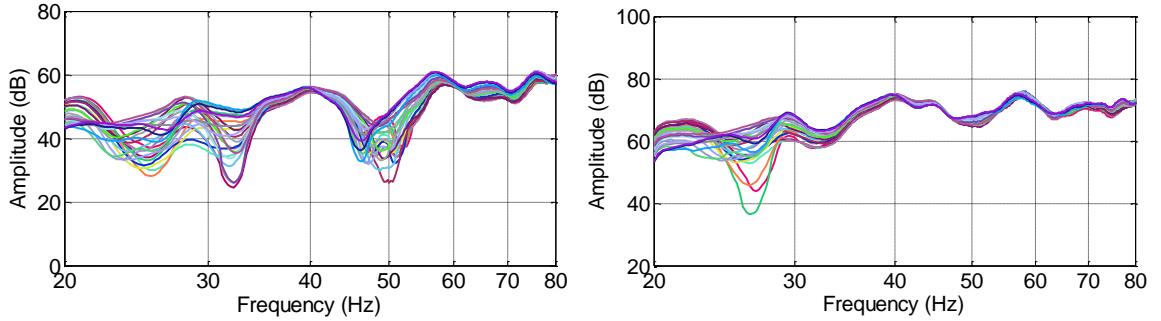


Fig. 5.51 Uncorrected (left) and corrected (right) hybrid subwoofer CSA measurements

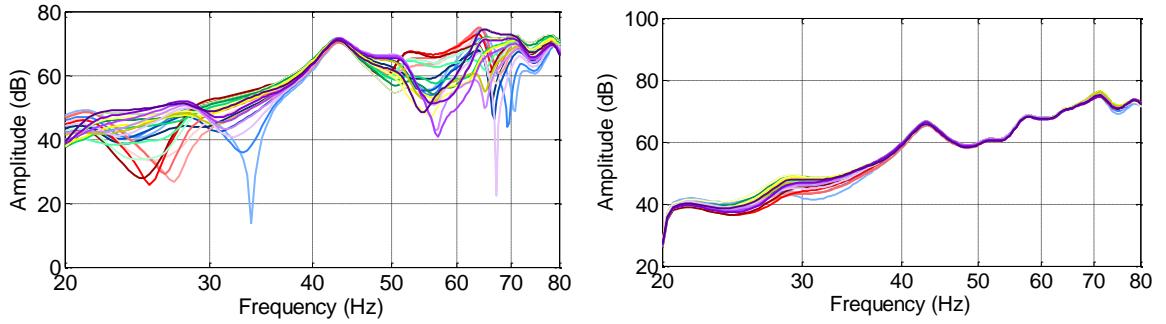


Fig. 5.52 Uncorrected (left) and corrected (right) hybrid subwoofer CSA simulations

The hybrid subwoofer CSA measured and simulated systems result in spatial variance reductions of 51.1% and 82.6%, respectively. The discrepancy in performance between the two lies in the hybrid subwoofer prototype. Unlike in the simulation, the prototype's degrees of freedom are not fully independent of one another. Since the enclosure is sealed, there is a small amount of cross-talk between units as the internal air volume is compressed or stretched. Crosstalk results in loss of source component orthogonality which is assumed within the correction algorithm, giving imperfect correction results (Fig. 5.51).

Since the dipoles are restricted at low frequency, both the simulation and measurements display issues below 35 Hz. In this case, the limit was set to just above 35 Hz, therefore only the

omnidirectional components are active, reducing the available degrees of freedom. The measured results do, however, indicate that a hybrid subwoofer CSA is viable and can provide spatial variance reduction on par with the omnidirectional subwoofer CSAs, but with a single enclosure as opposed to four.

5.6.4 Discussion

Two prototype CSAs have been evaluated in this section with direct comparisons to their simulated performance. With the exception of the individual subwoofer frequency responses, the measured and simulated four-unit omnidirectional subwoofer CSA results are in good agreement, giving strong indication that CSA processing can be practically applied to subwoofer systems as DSP. The measurements show around 60% spatial variance reduction, which is in the range of what is expected out of the CSA method.

The more interesting experiment in this section was with the single-unit hybrid subwoofer CSA (four degrees of freedom). This system is the type that has been theoretically and virtually developed throughout this chapter. As the tested system contains an equal number of degrees of freedom as the omnidirectional subwoofer CSA tested, the spatial variance results should be similar, assuming all source components are orthogonal.

The measured and simulated results in this case cannot be directly compared in terms of frequency response as the source components are not directly in agreement between systems. The spatial variance reduction is comparable, nonetheless, where the measured system gives around 50% reduction. Although there are clear cross-talk issues between the hybrid subwoofer source components, causing correction inaccuracies, as well as non-ideal polar patterns for the dipole source components due to the issues discussed in Section 5.1.2, the prototype system still shows that this approach to low-frequency correction is feasible.

Future work should address the cross-talk issues, possibly with an additional set of filters, so that each source component operates independently. Also, to improve practicality additional work must address low-frequency extension to compensate for the dipole inefficiencies. Addressing these two issues should make hybrid subwoofer CSAs a more robust system and potentially a marketable product.

5.7 Performance comparisons

Up to this point CSA low-frequency room response correction systems have only been compared to themselves. It is necessary to compare CSAs to the systems presented in Chapter 4 to best judge if a CSA is an improvement over existing methodologies. This section compares various CSA configurations to conventional correction methods including: passive placement (single and multiple subwoofer), parametric equalization, single-point static inverse filter equalization and multiple-point static average inverse filter equalization.

Simulations are conducted identically to those in the previous sections in a 5 m x 4 m x 3 m space. A single omnidirectional subwoofer in the room corner simulation serves as a reference (or benchmark) spatial variance to which all other configuration performances are compared. Seventeen configurations and correction methods are tested and spatial variance reduction percentages calculated to rate each method based on improvement from the reference system. Test results are shown in Table 5.4, with results ranked from best to worst. In the table, CSA system results are in green, conventional DSP results are in blue and passive layout results are in purple. Note that pseudo-inverse processing is not included in these comparisons.

#	Correction system type	Layout description	Spatial variance (dB)	Spatial variance reduction (%)
1	CSA	4 units, room corners	0.16821	96.4
2	CSA	4 units, wall midpoints	0.75496	83.8
3	CSA	4 units, two vertical stacks	0.76554	83.6
4	CSA	2 units, opposite room corners	0.86746	81.4
5	CSA	4 units, random placement	1.8356	60.7
6	CSA	4 units, across room front	2.0798	55.4
7	Manual 3-band PEQ	4 units, room corners	2.7787	40.4
8	Multi-point average EQ	4 units, room corners	2.7787	40.4
9	Single-point EQ	4 units, room corners	2.7787	40.4
10	Passive layout	4 units, room corners	2.7787	40.4
11	CSA	1 unit, room corner	3.1962	31.5
12	Passive layout	4 units, wall midpoints	3.2012	31.4
13	Manual 3-band PEQ	1 unit, room corner	4.6642	0.0
14	Multi-point average EQ	1 unit, room corner	4.6642	0.0
15	Single-point EQ	1 unit, room corner	4.6642	0.0
16	Reference system	1 unit, room corner	4.6642	N/A
17	Passive layout	1 unit, wall midpoint	5.0371	-8.0
18	Passive layout	4 units, across room front	5.1935	-11.3

Table 5.4 Low-frequency system performance results due to various configurations and correction methods (reference system = single subwoofer in a room corner)

Of the systems tested, the CSAs perform with the most consistency over the various layouts. The only CSA that fails to outperform the other methods is the single-unit CSA with room corner placement, due to its low number of degrees of freedom (four) and disadvantageous positioning (high modal coupling). The tested equalization methods confirm the point made in Chapter 6 that although these systems improve the response at a single location in the listening area, they do not affect spatial variance. All equalization methods in Table 5.4 perform the same as the corresponding passive layouts. Concerning the passive layouts, system configuration is of upmost importance since performance ranges from reductions of over 40% to *increases* in spatial variance of over 10%. The CSAs, on the other hand, generally give above 50% reduction regardless of the physical system layout.

The correction methods compared to the CSAs are not the most complicated and robust available in practice. More advanced methods, such as those discussed in Chapter 4.2, likely perform closer to the level of the CSAs. Unfortunately, these systems are either exceedingly advanced for straightforward implementation within the FDTD toolbox or their DSP algorithms are not publically available.

Nonetheless, the CSA low-frequency room correction methodology is shown to outperform the more common room-mode correction systems, both in terms of spatial variance reduction and insensitivity to physical layout. Testing gives further confirmation to the CSA layout recommendations presented in Section 5.4.5, where the systems with larger source spacing and more degrees of freedom give the best results.

5.8 Chapter summary

A low-frequency wide-area room-mode correction technique has been presented in the form of a chameleon subwoofer array (CSA). These arrays are ideally implemented using hybrid, multi-component subwoofers which allow for up to four degrees of freedom per unit as opposed to the single degree of freedom with traditional omnidirectional subwoofers. It is possible, however, to utilize omnidirectional subwoofers within a CSA, although this approach may require an unrealistically large number of units to achieve the desired spatiotemporal control over the listening area.

CSA physical configuration is not a trivial matter. Greater unit-to-unit spacing results in considerable sound pressure spatial variance reduction over a wide listening area without significant loss in system output efficiency. Care must be taken to avoid units in close proximity to room boundaries, as dipole units are ineffective in these locations due to early reflections from adjacent surfaces causing destructive interference at certain frequencies.

Overall, the longer the CSA (i.e. the more degrees of freedom) the better the system performance and response manipulation. In terms of practicality, an upper limit on CSA degree is set to four units, providing sixteen degrees of freedom. This allows for target points to be placed within one meter of one another, achieving accurate spatial sampling of the listening area. When target points are spaced far apart, points in between aren't likely to receive benefits from correction, therefore the system will only achieve isolated "pockets" of correction, not a uniform response over the entire area.

Pseudo-inverse filtering removes many of the target point constraints resulting from a restricted quantity of source components. This process allows for denser target point spacing, giving a better sampling of the listening area. Simulations and practical experiments have confirmed that this procedure is advantageous to the CSA algorithm and provides increased robustness to the system.

The in-depth development of CSA technology allowed for prototype testing; first with an omnidirectional subwoofer CSA and later with a single-unit hybrid subwoofer CSA. In both cases, spatial variance reduction of approximately 60% was achieved, which is in close agreement with simulations of similar systems. These experiments prove the feasibility of CSAs as effective low-frequency correction devices. Although further work is necessary to address certain issues present in the current construction of the hybrid subwoofer prototype, the CSA algorithm appears to translate well from the virtual world to the real world.

6 Virtual bass systems

Loudspeaker systems commonly operate with restricted low-frequency bandwidth due to design compromises, thus prohibiting the reproduction of the full audio bandwidth. One approach to this problem is based on the concept of the “missing fundamental” whereby a set of synthesized low-order harmonics causes a listener to infer the presence of the fundamental frequency even though it is not physically reproduced. Harmonic synthesis is achieved by applying nonlinear processing to the input signal in the form of a virtual bass algorithm to enhance the perception of low-frequencies. This methodology is of value to systems with small drive-units that cannot efficiently reproduce low-frequencies.

In the case of CSAs, however, low-frequency reproduction is not limited. Nevertheless, issues can manifest themselves in a few narrow frequency bands in which the CSA cannot perform efficient correction due to a combination of source-to-receiver coupling and target response requirements (as highlighted in Chapter 5). Instead of replacing the entire low-frequency band with virtual bass, the effect can be selectively utilized to ease correction requirements on the CSA (or any correction system, for that matter). As such, the process is much more subtle than normally encountered with virtual bass systems (VBS) which does much to disguise its somewhat artificial nature.

This chapter first discusses the principles of virtual bass; primarily the concept of the missing fundamental. Two approaches for VBS are highlighted, along with a discussion on their advantages and disadvantages. To overcome the disadvantages of each technique, a hybrid system was developed which is discussed in detail including subjective evaluation results. The hybrid virtual bass procedure can alternatively be implemented as a form of room-mode suppression, either as a standalone process or a supplemental tool within existing systems. This technique is highlighted, including specific remarks on implementation within the CSA architecture to alleviate the aforementioned issues concerning correction filter practicality. This discussion includes results from listening evaluations using a standalone version of this method to emphasize the concept’s potential usefulness in a wide range of sound reproduction systems.

6.1 Principles of virtual bass synthesis

The phenomenon of the missing fundamental, or the residue pitch, is a result of the complex pitch-extraction mechanism within the inner ear and brain. When presented with a spectrally-complex sound, the pitch extraction mechanism attempts to make sense of the received signal by relating various spectral components to one another [78]. Equally-spaced spectral components result in a perceived pitch corresponding to the greatest common factor of the frequency values (in Hz) that falls within the audible range of 20 Hz – 20 kHz. For instance, if the source contains spectral components at 200, 300, 400 and 500 Hz the overall perception corresponds to a harmonically-rich tone at 100 Hz. The same concept applies to more complicated harmonically-rich signals, where energy may already exist at the required harmonics. To achieve the virtual bass effect, new harmonic components must still be introduced with proper phase coherency in relation to the fundamental frequency. This requirement is discussed in detail in Section 6.3.

This effect can operate using only two higher harmonic components of the fundamental. Adding additional harmonics increases the sharpness of the signal timbre (sound quality) as the average frequency of the components increases [78]. When applying the missing fundamental within low-frequency applications it is important to keep the average frequency of all spectral

components to a minimum so that the perceived pitch is close in timbre to the fundamental. Minimizing the number of harmonic components also preserves the fidelity of the source signal since virtual bass components are a form of distortion which should ideally be kept to a minimum.

6.2 Time-domain implementation – nonlinear devices

A nonlinear device (NLD) is the most common harmonic generator used within VBS for a number of reasons. First, the NLD is memoryless, allowing for real-time applications. NLDs operate using a polynomial approximation of a chosen function. The calculated coefficients are applied to the input signal as defined in (6.1).

$$y = \sum_{i=0}^N h_i x^i \quad (6.1)$$

where, h is a vector containing the N polynomial coefficients with x and y representing the signal input and output, respectively [79].

NLDs operate in the time domain, applying virtual bass over all spectral components of the signal. This will introduce intermodulation distortion if there are two or more closely-spaced spectral components in the input signal. While it has been argued that these components cause minimal auditory artifacts due to psychoacoustical masking at the Basilar membrane in the inner ear [79], intermodulation distortion is an unwanted peripheral to the NLD virtual bass system, which must be handled with care.

Early NLD virtual bass research utilized a full-wave rectifier (FWR) [80]. The FWR is simple to implement, but suffers from the fact that it generates only even-order harmonics. A FWR applied to a 100 Hz pure tone results in harmonic distortion at 200, 400, 600 Hz and so on. Following the principle of the missing fundamental, this harmonic series should result in a perceived pitch of 200 Hz rather than 100 Hz. The perceived pitch is a full octave higher than the target pitch perception, giving an inaccurate virtual bass effect.

This problem has led to a significant body of research aimed to develop the ideal NLD for virtual bass. A wide range of NLDs are presented in [79], where they are each objectively and subjectively evaluated to best judge performance. The second exponential-type NLD in [79] rates highly in objective and subjective tests and was therefore chosen as the NLD for this work. The input-output relationship is shown in Fig. 6.1 (polynomial coefficients taken directly from [79] for the approximation of the exponential curve).

NLD VBS are implemented with a series of filters to facilitate approximate control of the spectral components. The input signal is first processed by a low-pass filter (LPF) with a cutoff frequency set to the upper limit of the required low-frequency extension. The low-passed signal is processed by the NLD, generating the harmonic components. Next, the NLD output passes through a bandpass filter (BPF) to remove the fundamental spectral components and to roughly shape the harmonic components. If a low-frequency boost is required (as opposed to bandwidth extension), the BPF can be replaced by a LPF to retain the fundamental components. After the BPF, gain is applied and the resulting signal is combined with a delayed version of the input signal. The NLD virtual bass process is displayed in Fig. 6.2.

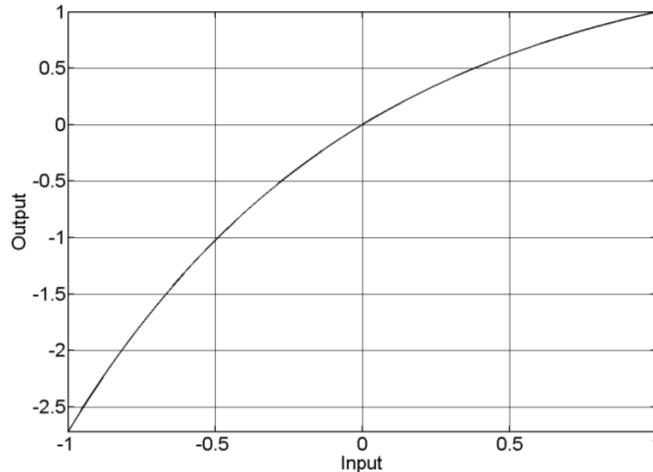


Fig. 6.1: Input-output relationship for the polynomial approximation of the exponential NLD

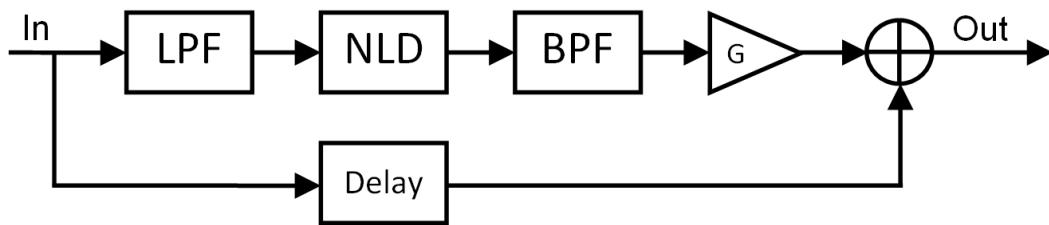


Fig. 6.2 NLD virtual bass procedure diagram

A widely-utilized commercial NLD-based VBS is called *MaxxBass* [81]. In addition to the system architecture in Fig. 6.2, *MaxxBass* uses equal-loudness processing to provide a virtual bass effect subjectively equal in level to the unprocessed signal. *MaxxBass* is available both in software and hardware form for use home theater systems as well as large-scale sound reinforcement applications.

6.2.1 Core possessing procedure and code

The first step required to apply NLD virtual bass to a signal is to synthesize and apply a low-pass filter (LPF) to the input signal (while retaining the original signal for later use). This particular implementation utilizes an FIR filter for straightforward delay calculations (Fig. 6.3). IIR filters could alternatively be utilized; however, IIRs are not linear phase which is undesirable for a high-fidelity sound reproduction system. After applying the filter, the resulting signal is normalized in order for the virtual bass algorithm to optimize nonlinear harmonic generation.

```

1 % generate the FIR LPF
2 N1 = 2000;                                     % filter order
3 SidelobeAtten = 40;                            % window sidelobe attenuation (dB)
4 win = chebwin(N1+1, SidelobeAtten);           % generate Chebychev window
5 b = fir1(N1, Fc/(Fs/2), ...);                 % generate filter coefficients
6 'low', win, 'scale');                          % generate digital filter
7 LPF = dfilt.dffir(b);                         % generate digital filter
8
9 % apply the LPF to the input
10 xL = filter(LPF, x);                         % apply filter to input
11 xL = xL/max(abs(xL));                        % normalize signal for processing
  
```

Fig. 6.3 MATLAB code for generating and applying a LPF to the input signal (x)

Once the signal has been processed by the LPF it is passed through the nonlinear device (Fig. 6.4). It is important to apply the LPF before virtual bass processing so that the nonlinear distortion components are generated only for the low-frequency components requiring replacement. Higher-frequencies should not be addressed within the procedure.

```

13      % run through NLD (see Eq. 8.1)
14      for n = 1 : 11
15          if n == 1
16              yL = h(n);
17          else
18              yL = yL + h(n) * (xL.^ (n - 1));
19          end
20      end

```

Fig. 6.4 MATLAB code for application of the NLD to the low-passed input signal (xL)

As per the NLD virtual bass procedure described in [79], the resulting signal must pass through a highpass filter (HPF) to remove the fundamental (low-frequency) components to achieve low-frequency extension through a band-limited loudspeaker, where reproduction at such low-frequencies is not practical (such as mobile device ear phones). The HPF is applied in a similar manner to the LPF (Fig. 6.5). If only bass boost is required, the HPF can be omitted.

```

22      % generate HPF1
23      N2 = 5000;                                % filter order
24      SidelobeAtten = 80;                      % window sidelobe attenuation (dB)
25      win = chebwin(N1+1, SidelobeAtten);       % generate Chebychev window
26      b = fir1(N2, Fc/(Fs/2), ...
27          'high', win, 'scale');                % generate filter coefficients
28      HPF1 = dfilt.dffir(b);                   % generate digital filter
29
30      % apply HPF1 to the processed low-frequency band
31      yL = filter(HPF1, yL);

```

Fig. 6.5 MATLAB code for generating and applying a HPF to the virtual bass signal (yL)

With the HPF applied, the system is ready for signal recombination. First, the original input signal must be processed with a HPF to remove the low-frequency components that were addressed in the VBS (Fig. 6.6).

```

33      % generate the FIR HPF2
34      b = fir1(N1, Fc/(Fs/2), ...
35          'high', win, 'scale');                % generate filter coefficients
36      HPF2 = dfilt.dffir(b);                  % generate digital filter
37
38      % apply HPF to the input
39      xH = filter(HPF2, x);

```

Fig. 6.6 MATLAB code for generating and applying a HPF to the input signal (x)

Before the signals can be recombined, they must be individually examined in order to match signal levels so that the two frequency bands are recombined with equal weighting. This is accomplished by inspecting the frequency response surrounding the crossover frequency and calculating the mean amplitude. The average amplitudes from each band are compared and a differential gain offset is determined (Fig. 6.7).

```

41 % determine appropriate gain offset
42 ref = 20e-6;
43 f1 = Fc*0.8;                                % lower frequency limit
44 f2 = Fc*1.2;                                % upper frequency limit
45 fVec = linspace(0, Fs, length(x));
46 ind1 = find(f1 < fVec, 1, 'first');          % frequency vector
47 ind2 = find(f2 < fVec, 1, 'first');          % lower index
48 X = fft(xH)/length(xH);                      % upper index
49 avgLX = 20*log10(...);                       % high-band frequency response
50     mean(abs(X(ind1 : ind2))/ref);           % average high-band frequency response
51 clear X;
52 Y = fft(yL)/length(yL);                      % low-band frequency response
53 avgLY = 20*log10(...);                       % average low-band frequency response
54     mean(abs(Y(ind1 : ind2))/ref);           % average response difference
55 clear Y;
56 diffL = avgLX - avgLY;                      % average response difference

```

Fig. 6.7 MATLAB code to find the differential gain offset between the two output signals

This gain differential is added to the required virtual bass gain (defined by the user for subjective calibration) and then converted to a linear scale and applied to the low-band signal (Fig. 6.8).

```

58 % apply gain to VB signal
59 GdB = GdB + diffL;                          % calculate low-band gain required
60 G = 10^(GdB/20);                           % convert gain to linear scale
61 yL = G*yL;                                 % apply gain to low-band

```

Fig. 6.8 MATLAB code to calculate and apply necessary gain to the low-band signal (yL)

Before the signals are recombined, delay must be applied to the high-band signal. Both signal bands have now passed through a FIR filter of order 2000 for crossover purposes. The low-band signal, however, has had a secondary filter applied to remove the fundamental low-frequency components; therefore the low-band signal is delayed from the high-band. This delay is calculated as half the FIR filter's order (filter order is always even within this specific routine) and applied to the high-band signal so that both signals are properly aligned in the time domain (Fig. 6.9).

```

63 % apply delay to input
64 temp = xH;                                  % store high-band signal
65 delay = N2/2;                               % calculate signal delay
66 xH = zeros(1, length(xH));                  % reinitialize high-band vector
67 xH(delay : end) = ...;
68 temp(1 : end - delay + 1);                  % apply delay to high-band signal

```

Fig. 6.9 MATLAB code to calculate and apply the delay to the high-band signal (xH)

Finally, the two signal bands are recombined to give the output signal which is band-limited to prevent attempted sound reproduction below the cutoff frequency of the radiating device with these components replaced by the virtual bass effect, implemented in this case by an NLD (Fig. 6.10).

```

70 % recombine the signals
71 y = xH + yL;

```

Fig. 6.10 MATLAB code to recombine the low- and high-band signals to form the output

6.2.2 Discussion

The primary advantage of NLD VBS is that the NLD is applied solely in the time domain. This lends itself well to real-time applications. Also, the procedure can result in very accurate transient performance as it does not require a tradeoff between time and frequency domain resolution. NLD approaches, therefore, are likely to perform particularly well for transient-heavy musical signals such as drum beats, staccato cello and double-bass.

However, a central disadvantage to NLD VBS is that it lends only approximate control over the frequency component content of the virtual bass effect. This control is achieved through the use of the HPF/BPF which shapes the frequency response of the effect. This control does not have the capability to address intermodulation distortion, since virtual bass is applied over all frequencies passed through the device and the harmonic content is dependent on the specific NLD utilized. The presence of intermodulation distortion can adversely affect signal fidelity and cause the effect to sound unnatural.

A further disadvantage of the NLD approach is that it is incapable of tracking pitched sounds. The procedure has no memory, therefore it treats each input sample as a discrete event. An NLD may see a pitched signal as a weak transient component, causing pitch to be poorly converted to the virtual bass domain.

Overall, NLD VBS are advantageous for real-time applications of virtual bass for transient-rich signals but are likely to suffer from reduced output for pitch signals such as vocals or organ. A frequency-domain comparison of input/output signals using this methodology is displayed in Fig. 6.11a, highlighting the harmonic-distortion components introduced to achieve the virtual bass effect while the low-frequency band has been attenuated based on the reproduction capabilities of the loudspeaker system (cutoff at 120 Hz, in this situation). The effect is also demonstrated as a bass boost application for the same input signal (Fig. 6.11b).

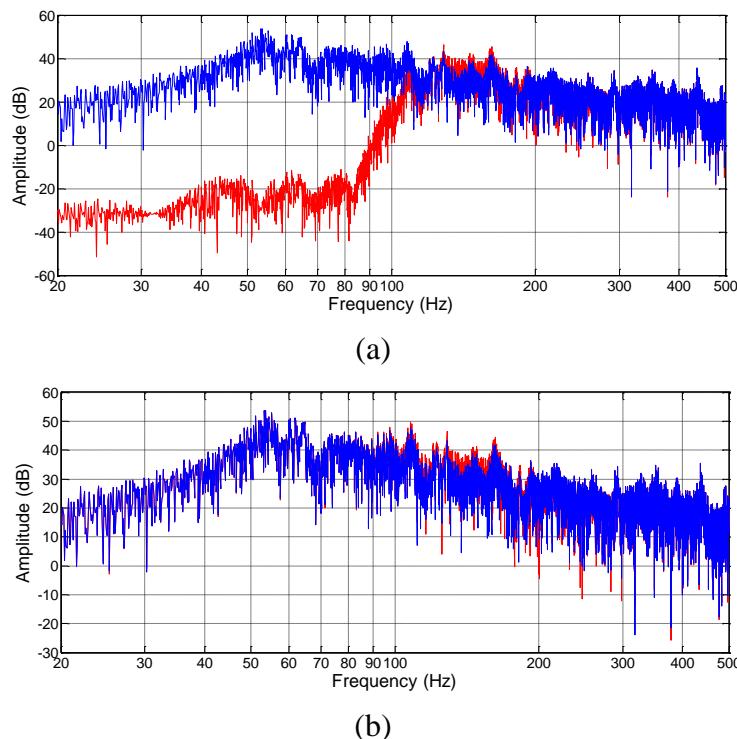


Fig. 6.11 Input (blue) and output (red) comparison using NLD
(a) low-frequency extension and (b) low-frequency boost

6.3 Frequency-domain implementation – phase vocoders

An alternative to NLD VBS has emerged in recent years utilizing a phase vocoder (PV) as the harmonic generator [82]. PV VBS provide superior harmonic control, allowing for selective harmonic inclusion. Since this approach operates in the frequency domain, intermodulation distortion can effectively be avoided, unlike with NLDs.

PVs operate by splitting an input signal into short time-domain windows (generally between 50 – 250 ms). The PV uses a fast Fourier transform (FFT) for each time window and applies the required processing while maintaining phase coherence and then generates the output signal either by sum-of-sinusoids or inverse Fourier transforms (IFFT) where each window is overlap-added to minimize amplitude-modulation effects. The present implementation utilizes sum-of-sinusoids.

A disadvantage of PV arises due to the trade-off between time and frequency resolution. PV VBS require adequate frequency resolution for accurate harmonic generation in addition to avoiding intermodulation distortion. Frequency resolution can be determined by (Eq. 6.2).

$$f_{res} = 1/t_w \quad (6.2)$$

where, f_{res} is the frequency resolution (Hz) and t_w is the window length (s). For example, a 125 ms window gives 8 Hz resolution while a 500 ms window gives 2 Hz. This issue can cause smeared transient performance which is evident when applied to signals such as drum beats.

Previous solutions have involved reinitializing the phase within the algorithm when a transient is encountered [83] and also removing any transients from the input signal and then reinserting them, unprocessed, at the PV output [84]. The phase re-initialization solution is difficult as it relies on precise transient detection; otherwise, phase re-initialization occurs in excess causing poor phase coherency for pitched signal components. The transient removal method has had low ranking in subjective tests since transient components are not addressed within the effect [84].

Even though the PV cannot handle transients perfectly it does perform well on pitched signal components. Unlike NLD systems, PV VBS do not require a LPF on the input stage, as the algorithm selectively applies the effect to frequency bins. Within the PV the selected frequencies are pitch shifted to the necessary harmonic frequencies and amplitude adjusted to match any equal-loudness requirements; therefore no BPF or HPF is necessary.

Since PV VBS are more computationally demanding, it is expedient to down-sample the input signal for real-time applications. This requires a LPF before the down-sampling process to avoid spectral aliasing. Once the signal has been processed, it is up-sampled to the original sampling rate and recombined with the delayed original signal. The PV virtual bass process is shown in Fig. 6.12.

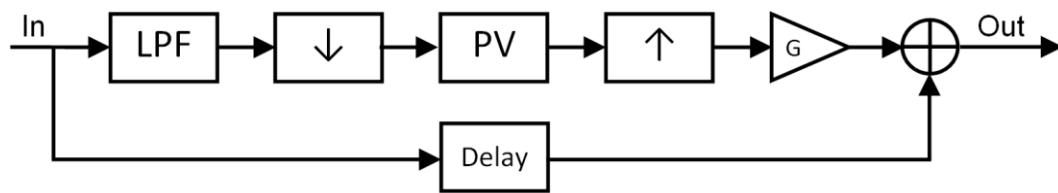


Fig. 6.12 Phase vocoder virtual bass procedure diagram

While PVs are commonly used for audio effects such as pitch shifting and time stretching [85], there are no known commercial applications of PV virtual bass.

6.3.1 Core processing procedure and code

The first steps in the PV virtual bass procedure store the initial input signal and sampling rate, define the new sample rate for the PV application and downsample the input signal (Fig. 6.13).

```

1      % store initial signal and sample rate for use later
2 -  xTot = x.';                                % store initial input signal
3 -  FsTot = Fs;                                % store initial sample rate
4
5      % copy and downsample signal for more efficient processing
6 -  Fs = 4096;                                 % reduced sampling rate
7 -  x = resample(xTot, Fs, FsTot);             % resample the input signal

```

Fig. 6.13 MATLAB code for input signal storage and downsampling procedure

Next, the cutoff for low-frequency extension is defined and the downsampled signal is passed through the corresponding LPF to remove irrelevant high-frequency components (Fig. 6.14).

```

9      % define target cutoff frequency
10 -  FcV = Fc;                                % PV cutoff frequency
11
12      % generate the LPF
13 -  N1 = 200;                                 % filter order
14 -  SidelobeAtten = 60;                        % sidelobe attenuation (dB)
15 -  win = chebwin(N1+1, SidelobeAtten);        % Chebychev window
16 -  b = fir1(N1, Fc/(Fs/2), ...                % generate filter coefficients
17      'low', win, 'scale');                     % generate digital filter
18 -  LPF = dfilt.dffir(b);
19
20      % apply LPF to the input
21 -  x = filter(LPF, x);

```

Fig. 6.14 MATLAB code for defining the VB cutoff frequency and applying the LPF

The PV parameters are initialized by defining the analysis window (w). In this case a raised cosine is used with a 250 ms length (N), with the window set to shift by 1/32 second each analysis step (HA). This ensures that there is sufficient frequency resolution (from the wide window) and also adequate transient handling (from the small window hop size). Once the window is constructed, the downsampled input signal is zero-padded so that the beginning and end of the signal are fully addressed within the VB effect. The necessary number of window steps for the particular signal is calculated ($pend$), to be used later in the processing loop and lastly, the output vector (yL) is initialized (Fig. 6.15).

Next, the specifics of the PV virtual bass effect are defined. This includes the range of harmonics to be generated, governed by the lower limit ($alphaL$) and the upper limit ($alphaH$). These are integer values, whereby a value of one corresponds to the fundamental. An integer step size ($step$) is defined to allow generation of even, odd or all harmonics. The frequency range is defined ($Fc1$ to $Fc2$) and is split into separate processing frames (nR frames, spaced at st Hz). The lower and upper limits for each range are placed into vectors (rL and rH , respectively) for use within the processing loop.

```

23      % generate the hanning window at appropriate length
24 -      N = 2^nextpow2(Fs/4);                      % window length
25 -      HA = N/8;                                  % hop size
26 -      w = (0.5*(1 - cos(2*pi*[0:N-1]/N))).';    % Hanning window
27
28      % add silence to beginning and end of input and normalize
29 -      L = length(x);                            % length of the input signal
30 -      x = [zeros(N, 1); x; zeros(N - mod(L, HA), 1)]...
31          /max(abs(x));                         % zero-padded + normalized signal
32
33      % determine number of windows needed
34 -      pend = length(x) - N;                     % analysis end index
35
36      % initializations
37 -      yL = zeros(L, 1);                          % output vector

```

Fig. 6.15 MATLAB code for PV window and output signal initialization

Separate processing frames are necessary since the lowest frequency bins' harmonics can still fall below the system cutoff frequency. This requires these frequencies to use higher harmonics generated for the virtual bass effect. The higher frequency frames do not require special handling since their harmonics will fall within the operating band. The processing code for this is shown in Fig. 6.16.

```

39      % set pitch shift parameter range (increments of 1, 1 = fund)
40 -      alphaL = pv1;                           % lowest harmonic generation
41 -      alphaH = pv2;                           % highest harmonic generation
42 -      step = pv3;                           % inter-harmonic step size
43
44      % set up frequency vector
45 -      fVec = linspace(0, Fs, N);            % frequency bin vector
46
47      % define the frequencies for target/non-target ranges
48 -      Fc1 = 20;                             % lowest processed frequency bin
49 -      Fc2 = FcV;                           % highest processed frequency bin
50 -      nR = 4;                               % number of processing frames
51 -      st = (Fc2 - Fc1)/nR;                % processing frame separation (Hz)
52 -      rL = Fc1 + st*linspace(0, nR - 1, nR); % low-limit target range vector
53 -      rH = Fc1 + st*linspace(1, nR, nR);    % upper limit target range vector

```

Fig. 6.16 MATLAB code for target harmonic definition and frequency frame initialization

The process enters a *for* loop to target each frequency bin frame independently. Before PV processing commences, the specific targeted frequency bins must be defined. This is accomplished using the lower/upper limit vectors defined in the previous code segment. The inclusive range of frequencies is used to populate the *indVec* which holds the index values of the targets bins. The non-targeted indices are placed within *otherVec*. These bins are still passed through the PV, but they are not pitch shifted. Lastly, a temporary output vector (internal to the loop) is initialized just before the PV algorithm begins (Fig. 6.17).

The PV virtual bass algorithm steps through each required pitch shift to achieve the desired harmonic components. This is performed within another *for* loop. Inside the loop, the Nyquist frequency (*L1*) is determined (based on the analysis window length, *N*) and the shift range is checked to ensure the generated harmonics are within the operating band. If not, the shift value is incremented by one in a *while* loop until the shifted range is appropriate. All PV components are next initialized (Fig. 6.18). This PV implementation is adopted from the work in [86].

```

55      % process each target range separately and sum the outputs
56 - for z = 1 : nR
57      % determine target/non-target frequency bins
58 -     indVec = find(rL(z) <= fVec, 1, 'first') : ...
59         find(rH(z) <= fVec, 1, 'first');% indeces of frequencies to shift
60 -     otherVec = [];% indeces of frequencies w/no shift
61 -     for n = 1 : length(fVec) % fill non-shift vector
62 -         if isempty(find(n == indVec, 1, 'first'))
63 -             otherVec = [otherVec, n];
64 -         end
65 -     end
66
67     % initialize temporary output vector
68 -     yT = zeros(length(x), 1);

```

Fig. 6.17 MATLAB code for target frequency bin definition and local output initialization

```

70      % run the procedure to generate the required harmonics
71 - for alpha = alphaL : step : alphaH
72 -     L1 = floor(N/2); % Nyquist evaluation length
73 -     shift = alpha; % pitch shift factor
74
75     % if frequency bins are being shifted to below the cutoff range,
76     % scale the shift value (integer) up until within proper range
77 -     while rH(z)*shift < FcV
78 -         shift = shift + 1;
79 -     end
80
81 -     oVec = ... % address freqs below Nyquist
82 -     otherVec(otherVec <= L1); % address freqs below Nyquist
83 -     omega = ... % sinusoid reconstruction vector
84 -     2*pi*HA*[0 : L1 - 1]'/N;
85 -     pin = 0; % initial measured phase
86 -     pout = 0; % initial magnitude
87 -     phi0 = zeros(L1, 1); % initial measured phase
88 -     r0 = zeros(L1, 1); % initial magnitude
89 -     psi = zeros(L1, 1); % initial phase
90 -     grain = zeros(N, 1); % analysis window vector
91 -     res = zeros(HA, 1); % resulting signal vector

```

Fig. 6.18 MATLAB code for PV routine variable checks and initializations

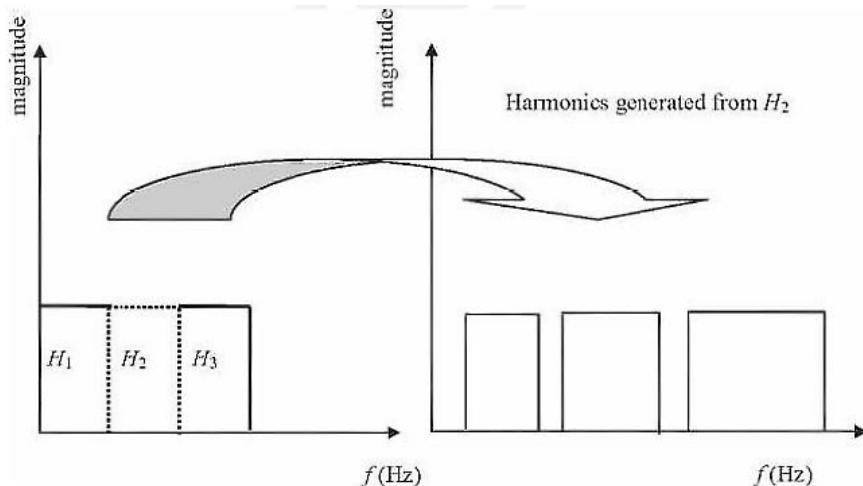
The PV virtual bass effect is applied by stepping through all analysis windows within a *while* loop. The crucial piece of code in this section is contained in a *for* loop from lines 106 – 115 (Fig. 6.19). The pitch shift is applied to the targeted frequency bins (*indVec*) during the phase adjustment stage by multiplying the phase increment (*delta_phi*) by the shift integer value (lines 111 – 112). This frequency shifting procedure is illustrated in Fig. 6.20 for clarity. The non-targeted bins (*oVec*) are also processed for phase adjustment, just without the shift multiplier (lines 109 – 110). The resulting complex frequency response is converted to the time domain (*res*) using a sum-of-sinusoids procedure (line 114) and is added to the temporary local output vector (*yT*) in lines 121 – 122. Once the PV virtual bass is applied, the temporary output is added to the overall output (*yL*) on line 132. This is the end of the core PV virtual bass procedure.

```

93 % run the procedure
94 - while pin < pend
95 -     grain = ...
96 -         x(pin + 1 : pin + N).*w;% analysis window
97 -         fc1 = fft(fftshift(grain)); % FFT of analysis window (w/circ shift)
98 -         f = fc1(1 : L1);           % analysis window up to Nyquist
99 -         r = abs(f);              % FFT magnitude
100 -        phi = angle(f);        % FFT phase
101 -        delta_phi = omega + mod((phi - phi0 - omega) ...
102 -            + pi, -2*pi) + pi;    % unwrapped phase diff.
103 -        delta_r = (r - r0)/HA;   % change in magnitude
104 -        delta_psi = delta_phi/HA; % phase increment
105 -        % apply linear interpolation over hop size (HA)
106 -        for k = 1 : HA
107 -            % shift each frequency bin
108 -            r0 = r0 + delta_r;      % adjust magnitude
109 -            psi(oVec) = psi(oVec) ...
110 -                + delta_psi(oVec); % adjust phase
111 -            psi(indVec) = psi(indVec) ...
112 -                + shift*...
113 -                delta_psi(indVec); % adjust phase w/shift
114 -            res(k) = r0'*sin(psi); % resulting signal (time domain)
115 -        end
116 -        phi0 = phi;              % save previous meas. phase
117 -        r0 = r;                  % save previous magnitude
118 -        psi = mod(psi + pi, -2*pi) ...
119 -            + pi;                 % unwrapped phase increment
120 -        % create time domain output
121 -        yT(pout + 1 : pout + HA) = ...
122 -            yT(pout + 1 : pout + HA) + res;
123 -        pin = pin + HA;
124 -        pout = pout + HA;
125 -    end
126 - end
127
128 % normalize output and cut out silence at beginning and end
129 - yT = real(yT(N/2 + HA + 1 : N/2 + HA + L));
130
131 % add to total output
132 - yL = yL + yT;
133 - end

```

Fig. 6.19 MATLAB code for the core PV virtual bass procedure (adopted from [86])

Fig. 6.20 PV harmonic generation procedure (processing for H_2 shown for a bass boost application where the fundamental frequency band is preserved) [82]

The remaining steps for PV virtual bass are similar to the NLD approach. First, the virtual bass signal is normalized and passed through an FIR filter which approximates an equal-loudness contour. Additionally, the signal (unlike with the NLD approach) is filtered by comparing input and output waveform envelopes. It is expected that the PV will smear the transient response of the signal due to the wide temporal analysis window. The output envelope is adjusted to match that of the input to give approximately similar transient characteristics as the unprocessed signal, except now with the virtual bass components included (lines 155 – 172, Fig. 6.21). Once filtering is complete, the signal is upsampled and is now ready for recombination with the high-band signal.

```

135      % normalize yL
136 -    yL = yL/max(abs(yL));
137
138      % create and apply inverse a-weighing filter
139      % (approximate equal-loudness contour)
140 -    N3 = 1000;                                % filter order
141 -    fVec2 = linspace(FcV, 1000, N3 - 7);     % frequency bin vector
142 -    m = aweight(fVec2);                      % magnitude weighting vector
143 -    f = zeros(1, N3);                        % filter frequency vector
144 -    f(1 : 6) = [0 Fc*0.1 Fc*0.2 Fc*0.3 ...
145           Fc*0.4 Fc*0.7]/(Fs/2);             % fill target frequency vector
146 -    f(7 : end - 1) = fVec2/(Fs/2);          % fill remaining frequencies
147 -    f(end) = 1;                            % ensure final frequency is 1
148 -    temp = m;
149 -    m = zeros(1, N3);                      % set freqs below cutoff to zero
150 -    m(7 : end - 1) = temp;
151 -    m(end) = 1;
152 -    b = fir2(N3, f, m);                  % generate filter coefficients
153 -    yL = filter(b, 1, yL);                % apply filter to VB signal
154
155      % determine envelope of input and apply to vb output
156 -    x = real(x(N/2 + HA + 1 : N/2 + HA + L)/max(abs(x)));
157 -    env1 = abs(hilbert(x));                 % input envelope
158 -    env2 = abs(hilbert(yL));                % output envelope
159 -    envD = env1./env2;                     % envelope ratio
160      % smooth envelope
161 -    tol = ceil(Fs/10);
162 -    for n = 1 : length(envD)
163 -        if (n + tol) > length(envD)
164 -            envD(n) = mean(envD(n - tol : length(envD)));
165 -        elseif (n - tol) < 1
166 -            envD(n) = mean(envD(1 : n + tol));
167 -        else
168 -            envD(n) = mean(envD(n - tol : n + tol));
169 -        end
170 -    end
171 -    envD = envD./max(abs(envD));           % normalize envelope
172 -    yL = yL.*envD(1 : length(yL));         % apply envelope to VB signal
173
174      % upsample processed signal
175 -    yL = resample(yL, FsTot, Fs);          % upsample VB signal

```

Fig. 6.21 MATLAB code for post-PV virtual bass signal filtering and upsampling

Next, the original input signal (stored at the very beginning of this process) must pass through a HPF to remove the low-frequency data (Fig. 6.22). As with the NLD approach, this step is unnecessary if a bass boost is required.

```

177      % generate HPF
178 -  N2 = 2000;                                % filter order
179 -  SidelobeAtten = 40;                         % sidelobe attenuation (dB)
180 -  win = chebwin(N2+1, SidelobeAtten);          % Chebychev window
181 -  b = fir1(N2, Fc/(FsTot/2), ...             % generate filter coefficients
182      'high', win, 'scale');                   % generate digital filter
183 -  HPF1 = dfilt.dffir(b);
184
185      % apply HPF to input
186 -  xH = filter(HPF1, xTot);

```

Fig. 6.22 MATLAB code for generation and application of a HPF to the input signal

Similar to the NLD virtual bass code, the two signal bands' frequency responses are compared to determine their gain offset, which is corrected to ensure the signals combine at comparable amplitude levels (Fig. 6.23).

```

188      % determine appropriate gain offset
189 -  ref = 20e-6;
190 -  f1 = FcV*1.2;                               % lower frequency limit
191 -  f2 = FcV*1.8;                               % upper frequency limit
192 -  fVec = linspace(0, FsTot, length(xTot));    % frequency vector
193 -  ind1 = find(f1 < fVec, 1, 'first');        % lower index
194 -  ind2 = find(f2 < fVec, 1, 'first');        % upper index
195 -  X = fft(xH)/length(xH);                    % high-band frequency response
196 -  avgLX = 20*log10...
197      mean(abs(X(ind1 : ind2)))/ref);           % average high-band freq response
198 -  clear X;
199 -  Y = fft(yL)/length(yL);                     % low-band frequency response
200 -  avgLY = 20*log10...
201      mean(abs(Y(ind1 : ind2)))/ref);           % average low-band freq response
202 -  clear Y;
203 -  diffL = avgLX - avgLY;                      % average response difference
204
205      % apply gain to VB signal
206 -  GdB = GdB + diffL;                          % calculate low-band gain required
207 -  G = 10^(GdB/20);                           % convert gain to linear scale
208 -  yL = G*yL;                                 % apply gain to low-band

```

Fig. 6.23 MATLAB code to calculate and apply necessary gain to the low-band signal (yL)

Delay is applied to the high-band signal, with care taken to ensure the delay is calculated properly since the low-band signal passes through filters at a lower sampling rate. Once the signals are temporally aligned, they can be recombined. If low-frequency extension is required, the final output signal is passed through the HPF used in Fig. 6.22 to remove any superfluous fundamental low-frequency data (Fig. 6.24).

```

210      % apply delay to HPF input
211 -  delay = ceil(((N1 + N3)*...
212      (FsTot/Fs) - N2)/2);                  % calculate signal delay
213 -  xH = [zeros(delay, 1); xH];            % apply delay to high-band signal
214
215      % sum the original and processed signals
216 -  if length(xH) > length(yL)
217 -      y = xH(1 : length(yL)) + yL;
218 -  else
219 -      y = xH + yL(1 : length(xH));
220 -  end
221
222      % apply HPF to output
223 -  y = filter(HPF1, y);

```

Fig. 6.24 MATLAB code for delay application and final signal combination and filtering

6.3.2 Discussion

The strengths of PV VBS contrast those of NLD VBS. NLD virtual bass is memoryless, lending itself well to real-time applications with exemplary transient handling but struggling to track steady, pitched sounds. The PV method, on the other hand, performs well applying the effect to pitched components due to the wide overlapping analysis windows. The downside to this approach is that the overlapping windows smear sharp transients in the signal. This has partially been addressed by the waveform window matching procedure shown in Fig. 6.21, but is not a complete fix to the problem, potentially allowing subjectively artificial sounding virtual bass for signals such as drum beats.

Another inherent downside to PV VBS is the computational requirements. The process is performed largely in the frequency domain, requiring numerous FFT operations along with nested loops to address all frequency bin frames and time-domain analysis windows. This problem is diminished by downsampling the input signal, but the PV still operates slower than the NLD. If adequate computational power is available, however, this may not prove crucial; nevertheless, it must be kept in mind when selecting a virtual bass procedure.

Overall, PV virtual bass is advantageous for steadier, pitched signals but is likely to suffer from smeared transients for signals such as drum beats. A frequency domain comparison of input/output signals using this methodology is displayed in Fig. 6.25a, highlighting the harmonic distortion components introduced to achieve the virtual bass effect while the low-frequency band is attenuated based on the reproduction capabilities of the loudspeaker system (cutoff at 120 Hz, in this situation). The effect is also demonstrated as a bass boost application for the same input signal (Fig. 6.25b).

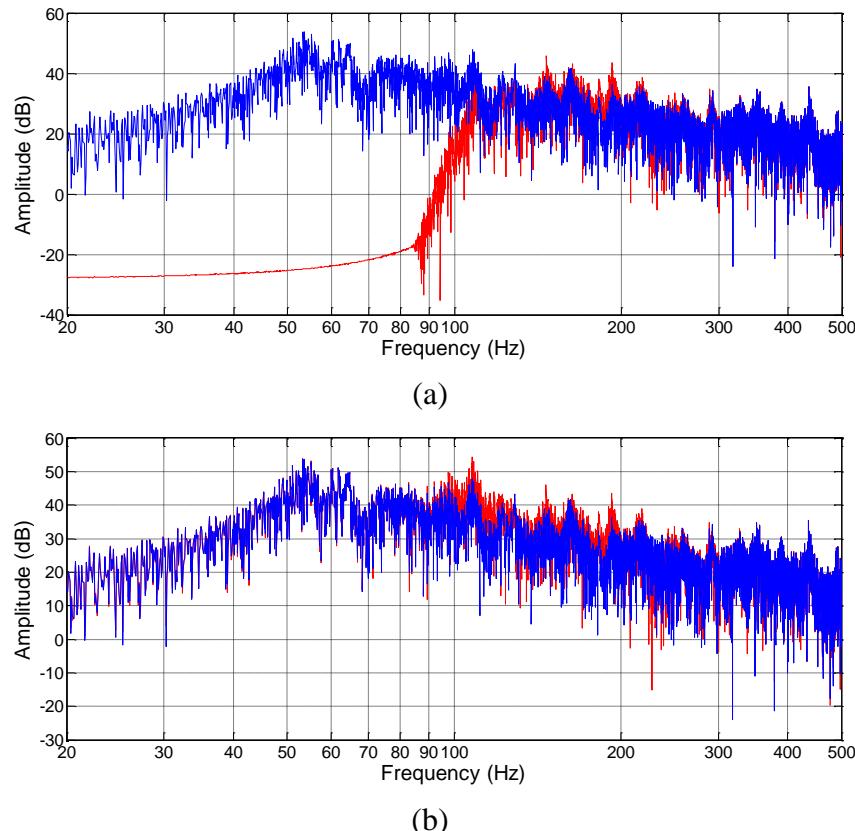


Fig. 6.25 Input (blue) and output (red) comparison using PV (a) low-frequency extension and (b) low-frequency boost

6.4 Hybrid implementation

A VBS that exploits the respective strengths of the NLD and PV systems but circumvents their weaknesses provides bass synthesis less sensitive to changes in input signal content. When the input signal has a high transient content, the system favors the NLD output and conversely, when the signal is more pitched the PV effect is utilized.

This hybrid approach requires a transient content detector (TCD) to analyze successive time domain windows of the input signal and appropriately weight the respective virtual bass algorithms running in parallel (Fig. 6.26).

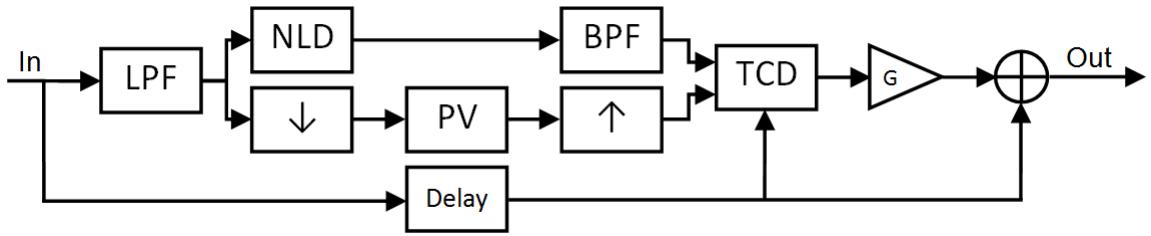


Fig. 6.26 Hybrid virtual bass procedure diagram

6.4.1 Transient content detector

The transient content detector (TCD) operates in the frequency domain, utilizing the improved low-frequency resolution of a constant-Q transform (CQT) [87]. Changes in spectral energy are tracked between successive analysis windows by targeting frequency bins in the virtual bass range (below ~ 100 Hz). When the change in energy exceeds a certain threshold, the overall weighting function is incremented in the NLD direction, while otherwise moving in the PV direction. This weighting function achieves the goal for the hybrid VBS where NLDs will be favored in transient-rich situations while PVs are utilized when the input signal contains steadier, pitched components. The TCD processing code is based on a non-real-time implementation where the entire signal is processed at the same time. This process could be built into the virtual bass code for real-time processing.

Like the PV processing, the TCD must normalize and zero-pad the input signal so that the algorithm addresses the beginning and end of the signal properly. Since the TCD uses a sliding analysis window (just like the PV), zero-padding eliminates the possibility of an abrupt beginning to the signal which would cause the weighting algorithm to fully favor the NLD. In order to focus on the low-frequency signal components, the signal is passed through a LPF and downsampled to improve processing efficiency (Fig. 6.27).

Next, the TCD analysis window is generated. The window characteristics are defined by the variables *fact* and *step*, which control the window length and hop size, respectively. In this implementation, the window length is set to 1/24 second with an analysis window hop size of 1/48 second. This ensures that even the sharpest transients are detected, including their onsets and decays. With the analysis window defined (indicating the number of windows required), an analysis window frequency response matrix (*inF*) is initialized along with vectors to store the mean spectral energy at each analysis window (*energy*) and the instantaneous rate of energy change between analysis windows (*dE*) (Fig. 6.28).

```

9      % normalize
10 -    x = x/max(abs(x));
11
12      % zero pad x
13 -    zpad = zeros(ceil(FsT/5), 1);
14 -    x = [zpad; x; zpad];
15
16      % apply LPF to the input
17 -    N1 = 1000;                      % filter order
18 -    Fc = fmax;                     % cutoff frequency
19 -    SidelobeAtten = 60;            % sidelobe attenuation (dB)
20 -    win = chebwin(N1+1, SidelobeAtten); % Chebychev window
21 -    b = fir1(N1, Fc/(FsT/2), ...
22         'low', win, 'scale');       % generate filter coefficients
23 -    LPF = dfilt.dffir(b);          % generate digital filter
24 -    x = filter(LPF, x);           % filter input signal
25 -    xo = x;                      % store resulting signal
26
27      % resample
28 -    Fs = 2^nextpow2(fmax*3);       % new sample rate
29 -    x = resample(x, Fs, FsT);

```

Fig. 6.27 MATLAB code for normalization, zero-padding, lowpass filtering and downsampling of the input signal for use within the TCD algorithm

```

31      % define window length and overlap
32 -    fact = 24;                      % length = 1/fact seconds
33 -    step = 2;                       % hop size = length/step seconds
34 -    L = ceil(Fs/fact);              % analysis window length
35 -    O = ceil(Fs/(fact*step));        % abalysis window hop size
36 -    w = tukeywin(L, 0);             % generate analysis window
37
38      % determine number of windows
39 -    nW = ceil(length(x)/O - step - 1);
40
41      % Create zero-padded x vector
42 -    x = [x; zeros(3*O, 1)];
43
44      % split signal into windows of FFT
45 -    inF = zeros(nW, L);             % matrix to store analysis windows
46 -    ind1 = 1;
47 -    ind2 = L;
48 -    energy = zeros(1, nW);          % vector to store avg spectral enegy
49 -    dE = zeros(1, nW);              % vector to store change in energy

```

Fig. 6.28 MATLAB code for TCD analysis window definition and matrix initialization

The TCD requires a time-frequency transform capable of sufficient low-frequency resolution to accurately identify the transient and pitched signal components. Discrete Fourier transforms (DFT) are commonly used for this purpose, but suffer from poor low-frequency resolution since the transform operates over linearly spaced frequency bins. This lack of spectral definition inhibits differentiation between transient and pitched signals in the low-frequency range (below 100 Hz); therefore it is necessary to utilize a transform which provides higher resolution.

DFTs maintain equal resolution through all linearly spaced frequency bins, but CQTs (commonly used in musical signal analysis) give frequency-dependent resolution based on logarithmically-spaced frequency bins [87]. This is achieved by scaling the time domain analysis window length based on the frequency bin under inspection (Eq. 6.3), allowing for CQT calculation through Eq. 6.4.

$$N_k = f_s Q / f_k \quad (6.3)$$

$$X_k = \frac{1}{N_k} \sum_{n=0}^{N_k-1} w_k(n) x(n) e^{-j2\pi Q n / N_k} \quad (6.4)$$

where, N_k is the analysis window length for frequency bin k , f_s is the system sampling rate (Hz), Q is the ratio of frequency to resolution, f_k is the frequency bin value (Hz) and X_k is the CQT for frequency bin k , with w_k representing the analysis window function applied to input signal, x [87]. The advantage of CQT over DFT is shown in Fig. 6.29 where a CQT is compared to a DFT with the same maximum window size for a signal consisting of pure tones from 20 – 100 Hz (10 Hz increments) and 100 – 1000 Hz (100 Hz increments).

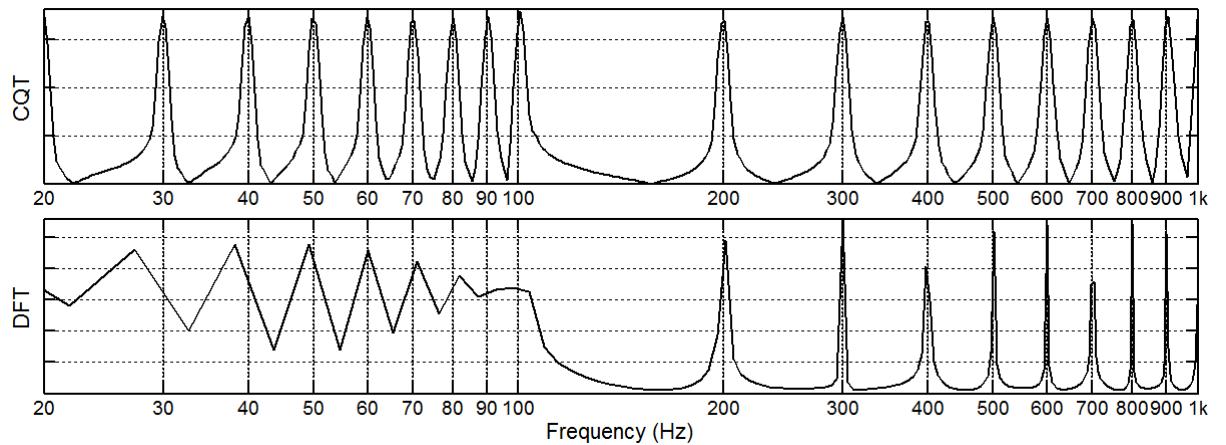


Fig. 6.29 CQT (top) and DFT (bottom) comparison for signal with sinusoidal components at each grid line (equal maximum analysis window size for both transform methods)

The CQT and DFT comparison highlights a tradeoff. Since the DFT has fixed analysis window size, the frequency resolution is equal (in Hz) over all frequency bins; therefore, on a logarithmic scale (as in Fig. 6.29) there is higher resolution as frequency increases, but below 100 Hz there are significant inaccuracies. The CQT method, while not as accurate at higher frequencies, maintains a constant frequency resolution to give equal accuracy on a logarithmic scale. This results in much more accurate spectral transforms for frequencies below 100 Hz due to the CQT's adaptive analysis window size. The improved accuracy of the CQT allows for detailed analysis within the TCD to best track the dynamic nature of an audio signal.

The CQT is implemented first by defining the desired Q-factor for the transform (Q). Next, the logarithmic frequency spacing factor (b) is calculated and utilized to determine the frequency bin locations (f_k). The maximum analysis window length is defined by the Q-factor and the lower frequency limit, which is used to determine the window lengths for each frequency bin. Lastly, all analysis windows are generated and stored in a matrix (wT) (Fig. 6.30). The system is now initialized to perform the CQT procedure.

The CQT procedure steps through the frequency bins one at a time, applying the analysis window to the time-domain signal and then performing a short time Fourier transform (Eq. 6.4) on the resulting data vector. The process is repeated for each analysis window storing the resulting frequency responses in the frequency response matrix (inF). The energy level at each analysis point is calculated and used to determine the change in energy between windows (as

long as the current analysis window is not the first). This value determines the virtual bass procedure weighting (Fig. 6.31).

```

51 | % setup CQT
52 - Q = 64;                                % target frequency resolution Q-factor
53 - b = L/log(2*fmax/fmin);                % calculate logarithmic spacing factor
54 - k = 0 : L - 1;                          % frequency bin index vector
55 - fk = fmin*(2.^((k./b)));              % logarithmically-spaced frequencies
56 - Nmax = floor(Q*(1/fmin)*Fs);          % maximum window length
57 - N = Nmax./(2.^((k./b)));              % window length vector
58 - alpha = 0.5;                            % vertical window offset
59 - wT = zeros(length(k), N(1));           % analysis window matrix
60
61     % generate frequency-dependent windows and place in matrix, wT
62 - for m = 1 : length(k)
63 -     n = 0 : (N(m) - 1);
64 -     temp = alpha + (1 - alpha)*cos(2*pi*n./N(m));
65 -     temp(ceil(length(temp)/2) : end) = 0;
66 -     wT(m, 1 : length(temp)) = temp;
67 - end

```

Fig. 6.30 MATLAB code for CQT parameter initialization

```

69 | % perform CQT procedure
70 - for n = 1 : nW
71 -     xT = [x(ind1 : ind2).*w; ...
72 -             zeros(Nmax - L, 1)];        % apply window to input + zero-pad
73 -     for m = 1 : length(k)
74 -         temp = 0;
75 -         for d = 0 : N(m) - 1
76 -             % calculate short time Fourier transform
77 -             temp = temp + xT(d + 1)*wT(m, d + 1)*exp(-i*2*pi*Q*d/N(m));
78 -         end
79 -         inF(n, :) = temp/N(m);      % save frequency response in matrix
80 -     end
81 -     ind1 = ind1 + O;
82 -     ind2 = ind2 + O;
83 -     energy(n) = ...
84 -         sum(abs(inF(n, :)))/L;    % calculate spectral energy
85 -     % determine change in energy if not the first analysis window
86 -     if n > 1
87 -         dE(n) = 1e3*(energy(n) - energy(n - 1));
88 -     end
89 - end

```

Fig. 6.31 MATLAB code for the CQT procedure and spectral energy calculations

The weighting vector is created by defining a threshold value for the change in energy to indicate a transient. The vector is stepped through, one window at a time, and assigned a 1 if the value exceeds the threshold and a 0 otherwise. The resulting weighting vector passes through a moving average filter to smooth the transition from transient to steady-state. This is required since informal subjective testing indicates that when weighting is digital in nature (1 or 0) the abrupt transition between virtual bass algorithms sounds unnatural. Smoothing the transitions eliminates this problem, allowing the system to be immune to one-off transients in otherwise steady-signals (Fig. 6.32).

```

91      % gate dE
92      maxDE = max(abs(dE));
93      thresh = 6;
94      mixC = zeros(1, length(xo));
95      con = ceil(FsT/Fs);
96      ind1 = 1;
97      ind2 = 1 + L*con - 1;
98      for n = 1 : nW
99          if abs(dE(n)) < (maxDE/thresh) % find steady parts
100             dE(n) = 0;
101         else % find transients
102             mixC(ind1 : ind2) = 1;
103             if (n > 1) && (dE(n - 1) == 0)
104                 mixC(ind1 - O*con : ind2 - O*con) = 1;
105             end
106         end
107         ind1 = ind1 + O*con;
108         ind2 = ind2 + O*con;
109     end
110
111     % create mixing curve between NLD and phase vocoder
112     % (+1 = full NLD, 0 = full PV)
113     mixC = mixC/max(abs(mixC)); % normalize weight vector (0 or 1)
114     step = ceil(FsT/5000); % define moving average step size
115     win = ceil(FsT/8); % define moving average width
116     ind1 = 1;
117     ind2 = 1 + win - 1;
118     while ind2 <= length(mixC) % smooth weight vector
119         mixC(ind1 : ind2) = mean(mixC(ind1 : ind2));
120         ind1 = ind1 + step;
121         ind2 = ind2 + step;
122     end
123
124     % cutout zero padding
125     mixC = mixC(length(zpad) + 1 : end);

```

Fig. 6.32 MATLAB code for algorithm weighting vector generation and smoothing

As an example, a dynamic time window was applied to a continuous 60 Hz sinusoid to test the performance of the TCD. The window was constructed to replicate alternating transient and pitched components. Ideally, the TCD should heavily weight the NLD when a transient is encountered and then quickly shift to the PV, while the TCD should remain at full PV weighting for the pitched signal components (Fig. 6.33).

The results in Fig. 6.33 highlight the TCD’s functionality as required for the hybrid VBS. The NLD is used for transient components while the PV is used for the steady, pitched parts. The smooth transition between operating states ensures no additional distortion will be introduced from the automatic mixing algorithm.

Although the TCD performs as expected with the synthesized test signals, it is helpful to perform tests on real-world musical signals to ensure the algorithm functions appropriately within the hybrid VBS. Two clips of music were chosen to test the TCD: one of a sparse drum beat (Fig. 6.34) and the other a slow, low-pitched vocal performance (Fig. 6.35).

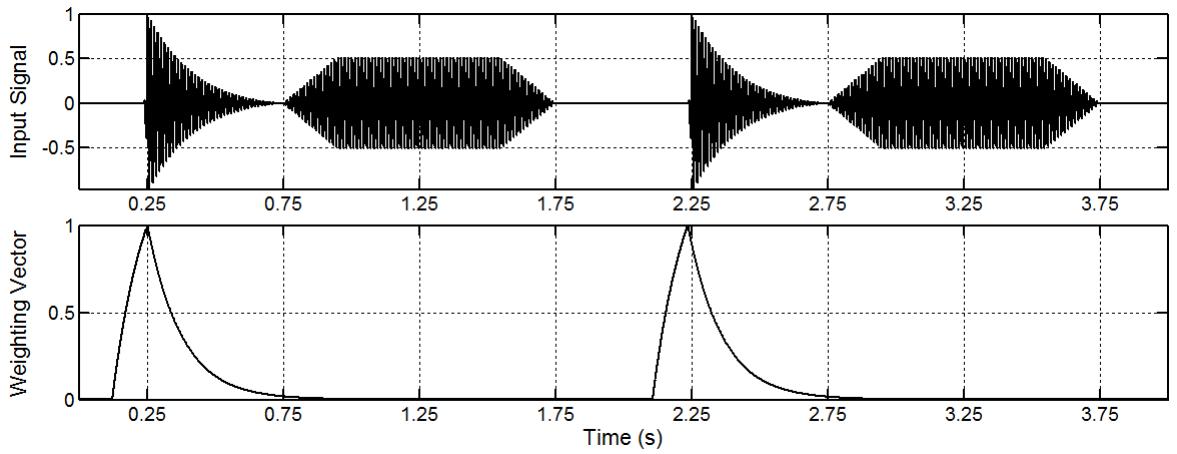


Fig. 6.33 TCD weighting function (bottom) due to input signal (top)
(0 = full PV weighting, 1 = full NLD weighting)

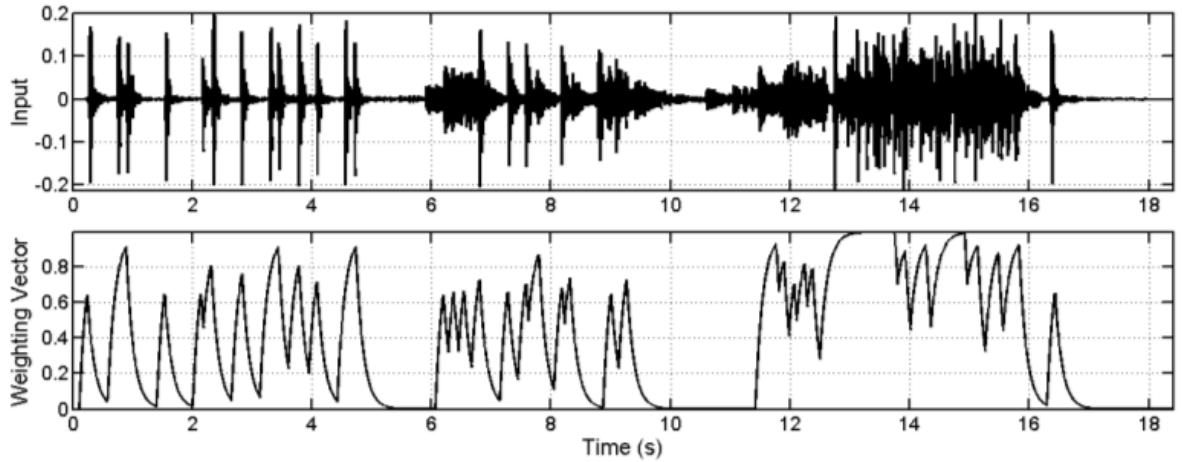


Fig. 6.34 TCD weighting function (bottom) due to drum beat sample (top)
(0 = full PV weighting, 1 = full NLD weighting)

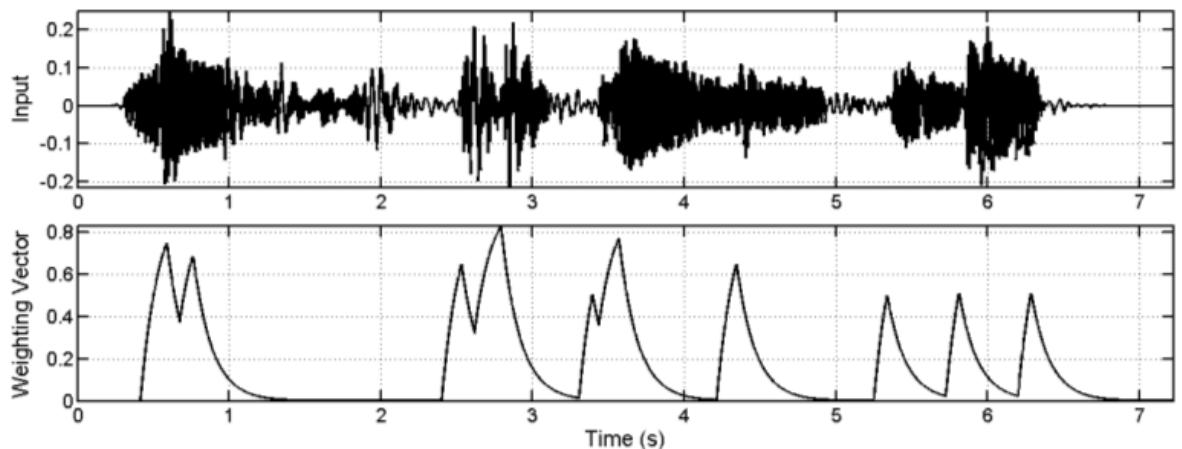


Fig. 6.35 TCD weighting function (bottom) due to vocals sample (top)
(0 = full PV weighting, 1 = full NLD weighting)

The TCD performs properly with the real-world music signals. The virtual bass effect for the drum beat consists almost entirely of the NLD output due to the transient-rich signal content.

The PV output is only mixed into the effect to enhance the resonances of the pitched drums after the initial transient attack. Similarly with the vocal sample, the NLD is only utilized for the initial onset of a note, while the PV handles the pitched components of the singing. The tests in Fig. 6.34 and Fig. 6.35 confirm the appropriate functionality of the TCD for use within the hybrid VBS.

6.4.2 Subjective evaluation

Subjective evaluation is required to evaluate the hybrid system in comparison to pure NLD and PV systems since objective measurements cannot provide insight to the virtual bass effect which occurs in the psychoacoustical domain. The hybrid system ideally should show less sensitivity to signal content (i.e. music genre and tempo) giving consistent ratings across a wide variety of musical stimuli, while the individual NLD and PV VBS should receive lower ratings for certain signal content due to their respective shortcomings.

The MUSHRA subjective testing method meets the needs of researchers in evaluating audio systems that cannot be classified objectively. The method, first developed within the BBC, was passed by the International Telecommunications Union (ITU) as ITU-R Recommendation BS.1534 [88]. The method requires subjects to be presented with multiple stimuli (MUS) where they are all available for listening at any time during the test. This allows subjects to compare the stimuli to determine their relative quality over an entire set.

Within each set of stimuli there must be a hidden reference and anchor (HRA). This provides a benchmark for the system under test. In the case of VBS, the reference is the unprocessed, full-range signal while the anchor is the high-passed version of the original signal with no virtual bass. In a properly constructed subjective test, the reference signal should receive the highest subjective rating while the anchor should get the lowest.

The rating scale ranges from 0 (bad sound quality) to 100 (excellent sound quality). Descriptors are placed at 25 point intervals (bad, poor, fair, good and excellent) to assist subjects to assign appropriate ratings.

A graphical user interface (GUI) was developed in MATLAB to meet the MUSHRA requirements (Fig. 6.36). The GUI presents subjects with ten stimuli, each containing different virtual bass enhancements applied to the same five-second music sample. The ten virtual bass samples are randomly placed within the sample space along with the reference and anchor.

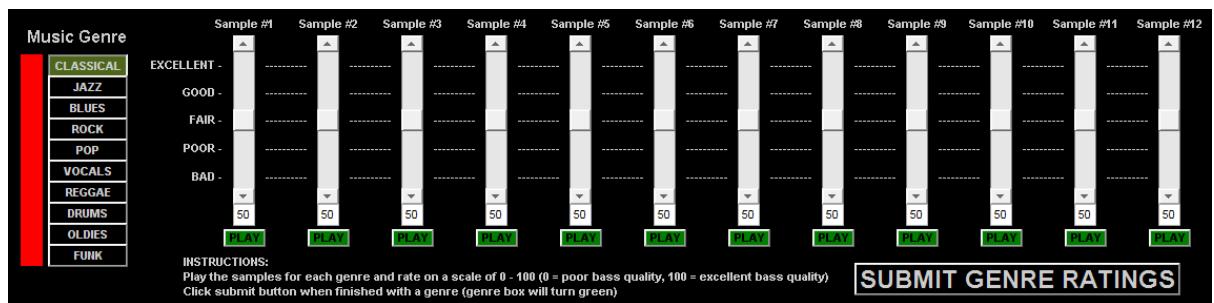


Fig. 6.36 GUI used for subjective virtual bass evaluation

Subjects proceed through ten genres of music (in any order) and rate the low-frequency sound quality of each sample. This allows for the VBS to be rated across a wide range of stimuli where certain ones will favor NLD or PV processing.

Tests were carried out within a quiet, isolated listening room with subjects left alone to complete the test. Subjects listened to the stimuli over a set of Beyerdynamic DT-770 headphones driven by a Sound Devices USBPre, connected directly to a PC. Sound levels could be adjusted to meet individual listening preferences.

While there was no time limit imposed on the subjects, it was recommended they spend approximately 3 – 5 minutes on each genre, resulting in a total test time of 30 – 50 minutes. Subjects were free to listen to stimuli as many times as necessary and could complete the genres in any order.

Twenty-one test subjects ranging in age from 23 to 63 (fifteen males and six females) participated. Although subjects found the test demanding in terms of duration, they all completed the required tasks, expressing overall enjoyment of the listening test.

The test results are presented in Fig. 6.37 where:

R = Reference (unprocessed signal)

A = Anchor (high-passed signal)

N = NLD system average

P = PV system average

H = Hybrid system average

with the results separated into their respective genres.

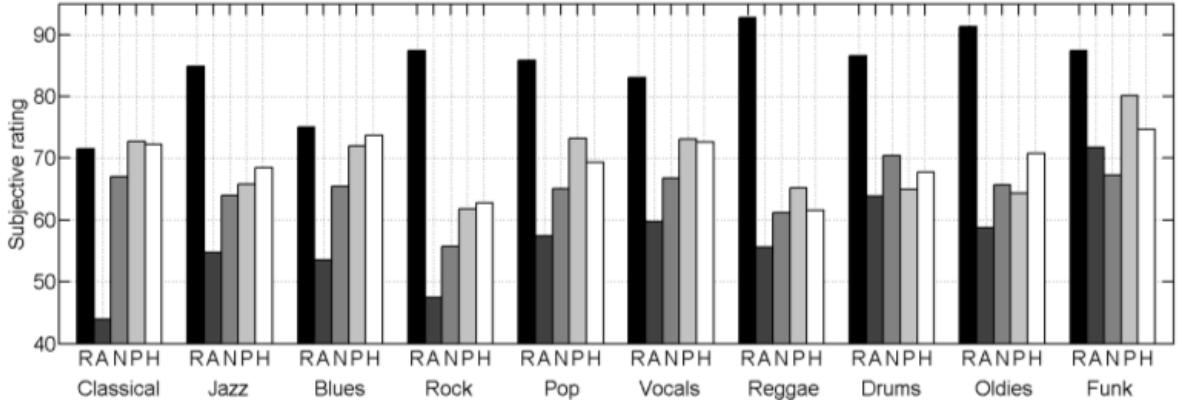


Fig. 6.37 Average subjective ratings for VBS
(R = reference, A = anchor, N = NLD, P = PV, H = hybrid)

The unprocessed (reference) sample achieved the highest ratings for all genres with the exception of classical. A lower classical rating can be because symphonic instruments produce harmonically-rich notes with weak fundamental components. While perceived pitch remains centered at the fundamental frequency, the higher harmonics are heavily weighted so the virtual bass effect may be perceived similarly to the unprocessed sample.

In all cases except funk, the high-passed (anchor) sample received the lowest ratings, although in a number of cases the anchor ratings were very close to some virtual bass ratings. These cases can be attributed to the VBS under question introducing a noticeably artificial quality to the sample.

The hybrid and PV systems outperformed the NLD for all genres except drums and oldies. These two genres are very transient-rich, lending themselves nicely to the NLD approach. In both these cases the hybrid approach outperformed the PV system, indicating that the TCD performed as expected, giving heavy weighting to the NLD output.

Although the hybrid VBS did not receive the highest rating for all genres, it never received the lowest. In most cases where it was not the best rated the hybrid system rated closely below the PV system or split the difference between the PV and NLD systems. This indicates that the hybrid system is generally less sensitive to input material, giving consistent performance.

6.4.3 Discussion

Hybrid VBS were developed not to achieve the best effect in every scenario, but rather to achieve consistent performance. NLDs and PVs show inconsistencies in performance when comparing transient-rich signals, such as drum beats, to steadier pitched signals, such as vocals or organ music. While the hybrid approach may not subjectively outperform the NLD or PV in every case, it shows consistent performance over all tested genres.

6.5 Room-mode correction

Virtual bass can be used as a supplemental component within a parametric equalization structure to help suppress just the most problematic room-modes without losing crucial audio information. When used as a standalone application, virtual bass often produces an artificial sounding effect which detracts from a natural listening experience. These applications are often targeted towards bandwidth extension of restricted loudspeakers where there are few alternatives to achieve low-frequency perception.

However, if virtual bass is limited to a narrow-band application, many of the artifacts can be masked by the surrounding frequencies of the physically reproduced energy. The narrow band(s) removed from the signal through parametric equalization can be reinforced psychoacoustically with the narrow-band virtual bass to maintain any information present at these frequencies. This room-mode correction approach is illustrated in Fig. 6.38.

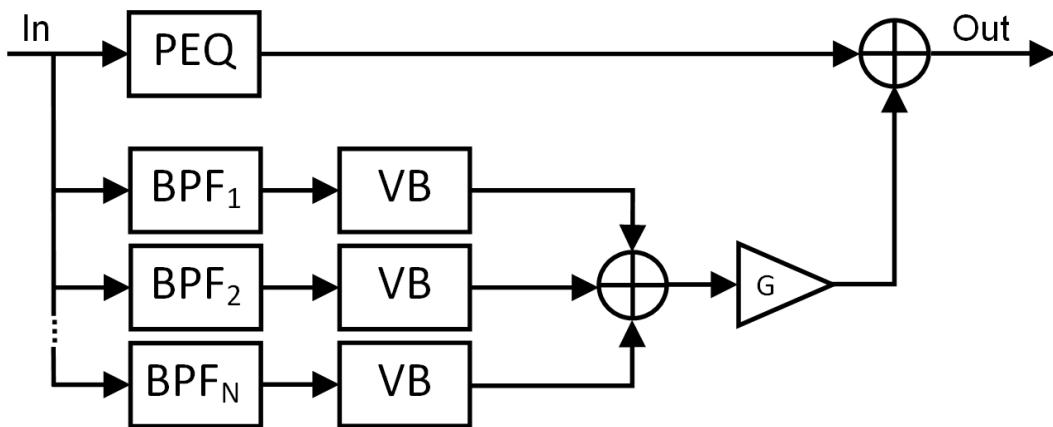


Fig. 6.38 Virtual bass room correction procedure

The unprocessed signal is sent via a parametric equalization (PEQ) routine with notch filters centered at the most problematic room-modes. The unprocessed signal travels in parallel through N bandpass filters (BPF_x) centered at each target (room-mode) frequency and is run through the hybrid virtual bass procedure (VB), as detailed in Section 6.4. All virtual bass outputs are summed with appropriate gain (G) applied to the resulting signal. The final virtual

bass signal is recombined with the parametric equalization output to give the fully processed signal.

This is not an entirely novel concept. A system is patented in [89] whereby virtual bass is utilized to avoid excessive “noise pollution” from a home theater system into neighboring residences. This methodology targets a frequency band spanning 50 – 120 Hz. The input signal is passed through a band-stop filter corresponding to the target frequency band and then, in parallel, the input signal has virtual bass applied to boost the impression of low-frequency energy. The two signals are lastly summed to form the output signal. Unlike the system in [89], the concept proposed in this work targets multiple narrow frequency bands in order to diminish spatial variance due to room-modes, as opposed to addressing “noise pollution” and/or band limited loudspeakers.

6.5.1 Subjective evaluation procedure

Since the virtual bass effect occurs within the human hearing mechanism and the brain, it is necessary to subjectively evaluate the proposed virtual bass correction procedure. Since the parametric equalization routine removes various narrow bands from the physically reproduced signal, the resulting signal is expected to have slightly reduced low-frequency impact. However, the procedure aims to maintain high-fidelity with minimal obvious artifacts due to the harmonic distortion from the virtual bass effect.

The evaluated sound reproduction system consists of two sealed-box subwoofers placed on the ground at wall midpoints to the left and right of the listening positions along with left and right main stereo loudspeakers. Two adjacent listening locations were chosen where the right location naturally receives strong low-frequency energy due to the close proximity of many mode antinodes while the left location receives little energy due to nodal placement. The four strongest room-modes (41, 58, 67 and 84 Hz) were chosen as targets based on FDTD simulations and confirmed with room measurements.

Ten high-fidelity musical recordings were used for the tests, each from a distinct musical genre as detailed in Table 6.1.

<i>Genre</i>	<i>Artist</i>	<i>Song</i>
Classical	Frank Zappa	<i>Dog Breath Variations</i>
Jazz	The Bad Plus	<i>Big Eater</i>
Blues	Bernard Allison	<i>Mean Town Blues</i>
Rock	Jeff Beck	<i>There's No Other Me</i>
Pop	Robert Randolph	<i>Diane</i>
Vocals	The Blind Boys of Alabama	<i>These Bones Gwine Rise Again</i>
Reggae	Bob Marley	<i>Get Up Stand Up</i>
Country	The Drive-By Truckers	<i>Bob</i>
Folk	Alison Breitman	<i>Tenafly</i>
Hard Rock	Audioslave	<i>The Worm</i>

Table 6.1 Musical selections by genre

Subjects were presented with an unprocessed musical sample and instructed to move between the two seats to judge both the overall sound quality and the low-frequency level variance between the two locations. Sound quality was rated on a one-hundred point scale with one-hundred being the best possible score. Low-frequency variance was also rated on a one-hundred

point scale with one-hundred representing significant spatial variance and zero representing no noticeable spatial variance.

The test subjects were given a list detailing the musical selections and asked to choose three based on their musical preferences. Next, a sample track was played, allowing subjects to become accustomed to the test procedure. Each unprocessed/processed clip pair was played until the subject had assigned ratings. The entire test required fifteen to twenty minutes to complete. The signal processing procedure was not revealed to the listeners to avoid any possible biasing.

6.5.2 Subjective evaluation results

The test subjects comprised of fifteen males and six females ranging in age from twenty-three to sixty-three years old. Each subject completed the test during independent sessions. The evaluation results are presented in Fig. 6.39, where the uncorrected/corrected seat-to-seat variance ratings (the two rightmost bars in Fig. 6.39) are the crucial factors in this investigation as they represent the perceived difference in frequency response between the two seats.

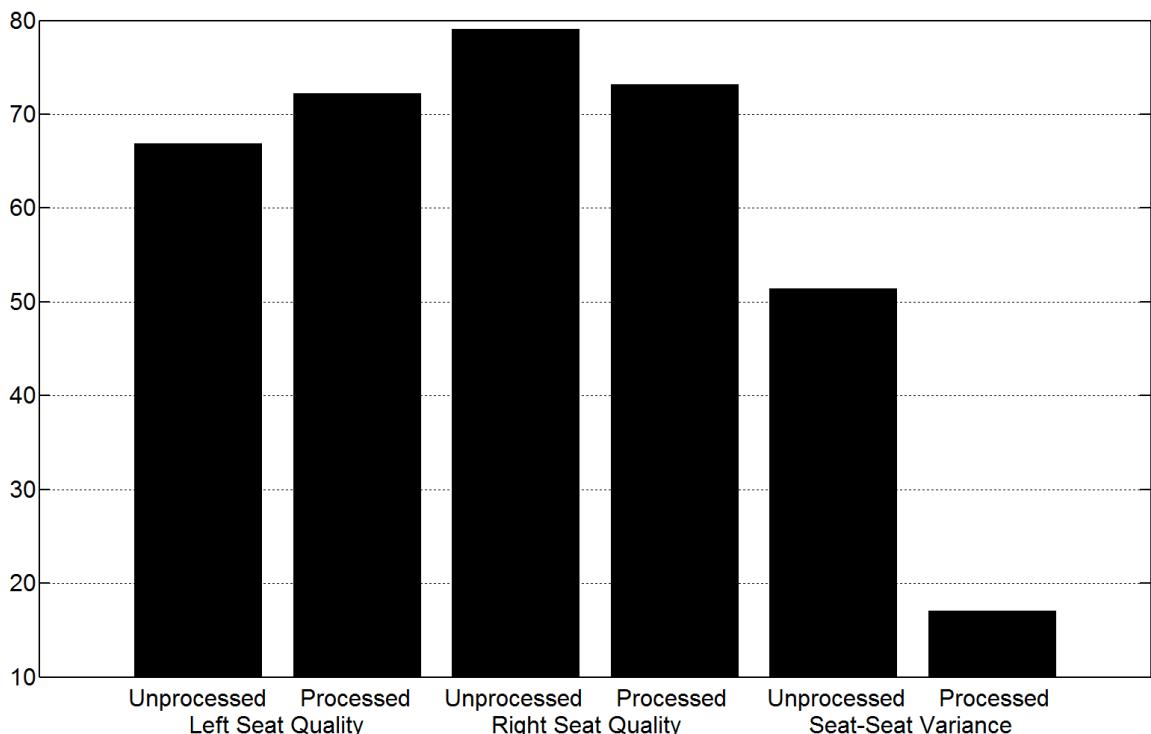


Fig. 6.39 Virtual bass low-frequency room-mode correction subjective evaluation results

The results indicate that the right seat receives quality ratings in the “good” range which can be attributed to a strong low-frequency presence. The left seat, on the other hand, receives quality ratings in the “fair” range with subjects commenting that they sensed the seat lacked certain musical information. The differences in quality ratings are reflected in the unprocessed seat-to-seat low-frequency spatial variance ratings in the “moderate” range.

After virtual bass processing, though, the subjective ratings show a noticeable shift. The right seat, while rated “good” unprocessed, has decreased to the lower bound of the “good” range. The left seat received “fair” quality ratings unprocessed, but has increased into the “good” range. The left and right seats’ processed ratings are within two points of each other which is

strongly reflected in the processed seat-to-seat low-frequency spatial variance ratings in the upper portion of the “not noticeable” range.

The subjective evaluations show that virtual bass room-mode correction provides a reasonable amount of spatial variance reduction. The compromise is that seats with naturally superb responses tend to experience slight decreases in fidelity in order to increase the fidelity of naturally poor sounding seats. Virtual bass ensures that all musical information present in the unprocessed signals is perceptually maintained.

6.5.3 Chameleon subwoofer array applications

The virtual bass room-mode correction procedure can be utilized as a supplemental component to any correction methodology. It was developed, however, with CSA applications in mind. As demonstrated in Chapter 5, the physical array/listener configuration and the target frequency response(s) may cause unreasonably high amplitudes across narrow frequency bands in the CSA correction coefficients. This causes the system’s mean output level (MOL) to drop considerably, due to the problematic frequency bands. Applying virtual bass correction to the CSA algorithm removes the problematic frequency bands from correction to keep filter coefficients at acceptable amplitudes.

The narrow bands targeted within the virtual bass procedure must be automatically detected within the CSA processing routine. This is accomplished using two metrics in tandem. First, the condition number of the measurement matrix is computed for each frequency bin. The condition number is used to indicate the sensitivity of a matrix to small system variations. This is critical in this case as system stability is essential for the practicality of CSAs, as discussed in Chapter 5. A low condition number indicates a well-conditioned matrix, while extremely high condition numbers highlight ill-conditioned matrices. The second collection of data are the CSA correction coefficients, themselves. The coefficient sets for the source components are averaged over the frequency spectrum for the overall required correction amplitude at each frequency bin.

Target band identification is performed by normalizing each metric to unity and then multiplying the two so that resulting values at or near unity represent particularly problematic frequencies. Here this value is referred to as the *instability factor*. A threshold is set, whereby any frequencies with instability factors above the threshold prompt exclusion from physical correction and replacement with virtual bass components. The threshold utilized for this exploration is the mean value over the intensity factor vector, although future work may focus on developing an improved threshold.

This technique for improved CSA functionality (in terms of maintaining acceptable MOL between uncorrected and corrected systems) is tested using six different system configurations in a 5 m x 4 m x 3 m room with 20 cm grid spacing and 10% absorption, all within the FDTD simulation toolbox. The configurations are as follows:

- 1) One hybrid subwoofer, room corner
- 2) Two hybrid subwoofers, opposite room corners
- 3) Four hybrid subwoofers, room corners
- 4) Four hybrid subwoofers, wall midpoints
- 5) Four hybrid subwoofers, room front
- 6) Four hybrid subwoofers, two vertical stacks $\frac{1}{4}$ along the side walls

The systems were first simulated uncorrected, noting the MOL, and then simulated after CSA correction, once spanning all subwoofer band frequencies and then again with the identified frequencies deleted from correction (example shown in Figure 6.40). The percent change in MOL is utilized as an indicator of the procedure's performance (Table 6.2).

#	Uncorrected MOL (dB)	Standard CSA Corrected MOL (dB)	Standard CSA Corrected MOL (% change)	VB CSA Corrected MOL (dB)	VB Corrected MOL (% change)	MOL % change due to VB processing
1	65	70	+7.7	80	+23.1	+15.4
2	85	75	-11.8	85	0.0	+11.8
3	90	90	0.0	95	+5.6	+5.6
4	85	75	-11.8	85	0.0	+11.8
5	90	45	-43.8	60	-25.0	+18.8
6	90	75	-16.7	85	-5.6	+11.1

Table 6.2 Simulated results comparing CSA correction with and without virtual bass (VB) replacement of problematic frequencies (MOL = mean output level)

The simulation results of the virtual bass replacement of problematic room modes give a clear indication that this method provides greater MOL for each system tested (generally around 10%). This eases the requirements on the physical sound reproduction system, granting more headroom for higher output levels.

One important characteristic is that the transitions between the included and excluded frequency bands (as seen in Figure 6.40b) are not instantaneous. To avoid the introduction of filter ringing due to what is commonly referred to as a “brick wall” filter, the transitions were smoothed by applying a window to the excluded bands. This eliminates sudden jumps from non-zero filter coefficients to zeroed coefficients and provides a more acceptable filter response.

The process highlighted in this section, although specifically implemented within the CSA algorithm, can be similarly applied to any low-frequency room-mode correction system (many of which are discussed in Chapter 4). Regardless of the correction system the virtual bass procedure is applied to, this method improves system output efficiency post-correction while not losing any aural data by subjectively replacing the physically removed frequencies with virtual bass components. As with the other proposals targeted at improving CSA practicality (presented in Chapter 5), this technique further strengthens the argument that CSAs can be realistically implemented without huge concern for practical nonlinearities.

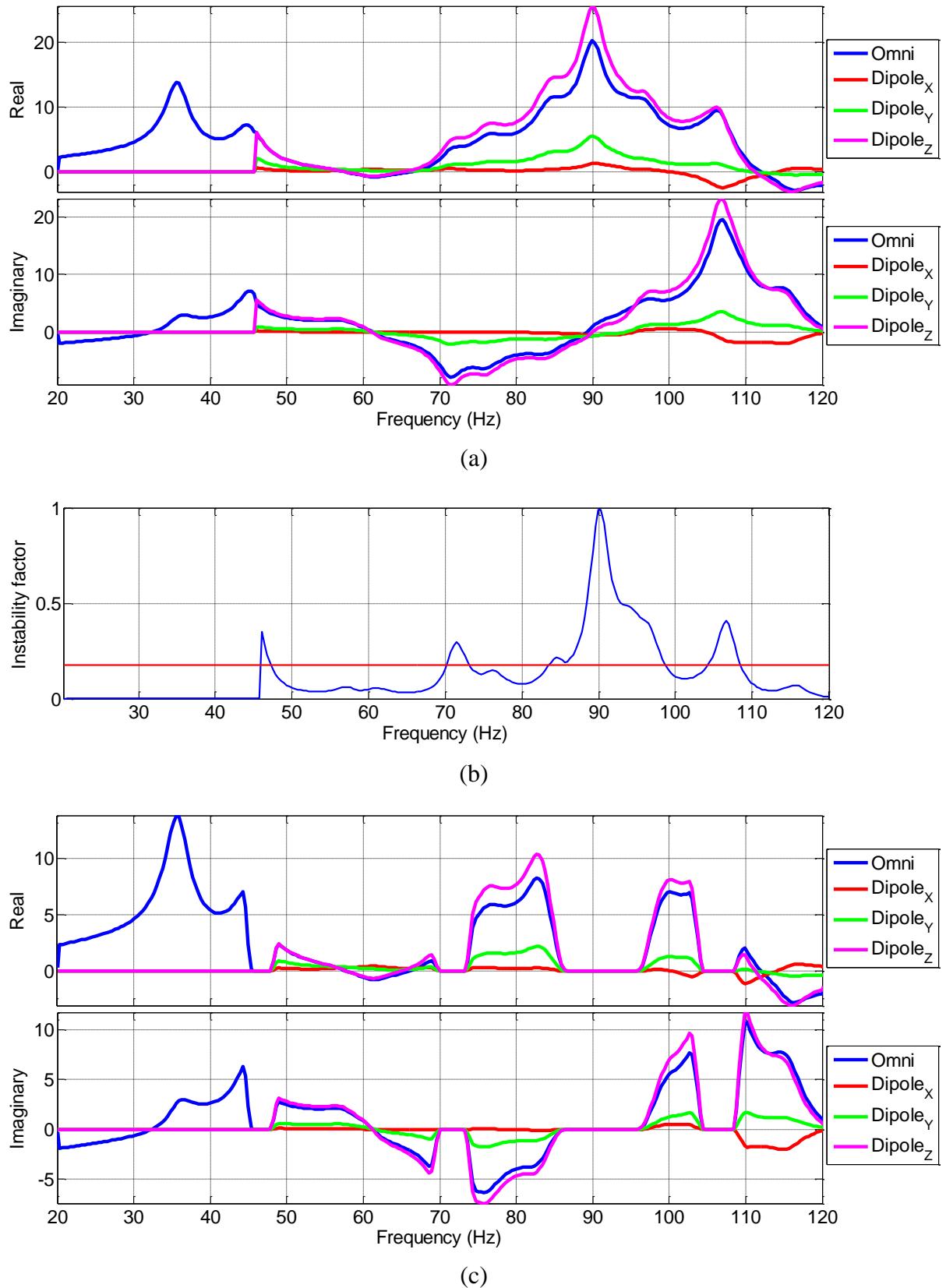


Figure 6.40 Example of problematic frequency band identification and removal for CSA processing including: (a) initial CSA correction coefficients, (b) instability factor values (red line = threshold) and (c) CSA corrected coefficients with problematic bands removed

6.5.4 Discussion

The impetus behind the development of virtual bass room-mode correction is to ease system requirements in existing methodologies where control at select narrow frequency bands proves elusive. Although subjective evaluation of the standalone virtual bass correction strongly indicates that spatial variance has been greatly reduced between listening points, test subjects expressed that the corrected system lacks a certain physical impact, which is due to the absence of physical reinforcement at the targeted frequency bands. In addition to reduced physical impact, overuse of virtual bass correction can result in perceptually unrealistic sound reproduction due to the harmonic distortion introduced. These two issues support the proposal that this form of correction be utilized as a supplement, rather than a replacement, to a well-established room-mode correction system.

As highlighted in Section 6.5.3, CSA correction may struggle at a few spectrally narrow regions. If the virtual bass procedure were implemented to handle subjective reinforcement over these frequencies, there would be minimal loss in physical impact while introducing nominal harmonic distortion from the virtual bass algorithm. The limited use of virtual bass is crucial as excessive harmonic distortion leads to a perceived “harshness” or “roughness” in musical signals [90], which is undesirable in high-fidelity sound reproduction. Introducing a mixing control of sorts could provide greater control of the virtual bass components by allowing users to select the attenuation required for the problematic frequencies and correspondingly mix in the virtual bass components. This concept can be designated for future work concerning this area.

6.6 Chapter summary

This chapter focuses on the possibility of using virtual bass as a low-frequency room-mode correction procedure. This is particularly valuable as a supplement within existing correction systems, including CSAs. Narrow frequency bands that a system struggles to correct efficiently can be removed from the physical reproduction and replaced by virtual bass components, thus preserving all aural information from the input signal.

The two methods for nonlinear bass synthesis, nonlinear devices (NLD) and phase vocoders (PV), each exhibit their own unique drawbacks. NLDs are memoryless to give excellent transient handling but struggle at detecting and processing pitched components and can produce intermodulation distortion. PVs utilize a sliding analysis window and process each time-domain window in the frequency domain. This approach allows for precise harmonic component manipulation, where the NLD applies harmonics based on the specific nonlinear input-output relationship chosen. While this windowed approach allows for proper handling of pitched signals, the windowing procedure smears sharp transients, thus diminishing the PV virtual bass subjective performance.

Combining the two approaches into a hybrid virtual bass system (VBS) is demonstrated in this chapter as a solution to the shortcomings of the two approaches. The hybrid system contains a dynamic weighting function between the approaches which operates under the control of a transient content detector (TCD). The TCD analyzes the input signal using a sliding analysis window by measuring the instantaneous rate of energy change between successive windows. If the energy change surpasses a pre-defined threshold, the hybrid system applies greater weighting to the NLD, which performs better with transient-rich signals. Conversely, if there is minimal transient content detected, then the system favors the PV, thus taking advantages of its pitched signal superiority. Subjective tests confirm that the hybrid system works consistently

well over many different musical genres, where the NLD and PV are shown to struggle with certain genres.

The hybrid system is utilized within a standalone virtual bass room-mode correction technique, which was implemented for use with subjective testing to judge if the procedure is practical for spatial variance minimization without losing signal information or fidelity. Subjective tests confirm that the approach works as planned, whereby two adjacent listening locations which naturally receive very different frequency responses due to strong room-modes were corrected to receive nearly equal subjective ratings without significant variation between the two seats.

Test subjects commented that the corrected system lacked some low-frequency impact that was present in the uncorrected system. This point emphasizes the suggestion that the virtual bass room-mode correction procedure be utilized only as a supplementary system within existing correction systems. Potential applications within the CSA architecture have shown an increase in mean output level of up to around 20%, thus reducing the requirements of the sound reproduction system.

Virtual bass can be a powerful aid within room-correction procedures. Caution must be taken, however, to restrict its use to only frequency bands requiring extreme correction since the effect purposely introduces harmonic distortion, which should be kept to a minimum to preserve system fidelity. If implemented carefully, though, a virtual bass supplemental procedure can greatly increase system performance and robustness.

7 Extended applications of chameleon subwoofer arrays

The primary application of chameleon subwoofer arrays (CSA) is to provide consistent spatiotemporal sound reproduction over a wide listening area in a small room. All locations within the listening area are targeted to exhibit identical responses after correction. This procedure is described in detail in Chapter 5 with a supplemental system component proposed in Chapter 6.

The CSA architecture lends itself to extended applications outside of spatial variance minimization. These alternative methods are explored in this chapter and range from providing personalized low-frequency control over a group of listeners to meeting both audience and musician sound reinforcement requirements in live sound settings.

In some cases, these alternative implementations are shown to be more practical than the primary CSA application, as they do not require additional hardware (the procedure exists entirely within the digital signal processing domain). Each proposed application is analyzed in terms of its ability to meet the low-frequency control specifications as well as its practicality for implementation within a commonplace sound reproduction system.

7.1 Location-specific equalization

Numerous methodologies for low-frequency control over a wide listening area exist; many of which are highlighted in Chapter 4. The vast majority (if not all) of these techniques aim to deliver either an even frequency response at all listening areas or to optimize the response at a single location. Few of these systems consider the influence of listener subjectivity which can vary considerably over a group of individuals. A system allowing each user to independently tailor the low-frequency response without affecting other listeners could accommodate subjective preferences and facilitate the listening experience.

The multiple degrees of freedom available within the CSA construct allow for each target location to be individually controlled, permitting multiple distinct frequency responses to exist in space simultaneously. This section explores the feasibility of this approach within small-sized listening rooms, similar to those used for analysis in Chapter 5.

7.1.1 Low-frequency control procedure

The individualized low-frequency control procedure is incorporated within the CSA correction filter algorithm and is presented here as an extension to the GUI used to calibrate the system. The system processes all measurements to determine the initial target frequency responses (room average response) at the listening locations. The user is then queried whether location-specific equalization is required. If yes, a separate GUI activates to guide the user through the individual manipulation (Fig. 7.1).

The GUI initially displays all target responses as identical (the default setting for CSA correction). The user selects any listening location to adjust. The adjustments are performed using a parametric equalization approach whereby the user has control over the center frequency, Q and gain of each filter. When a user is content with a specific filter, the “apply” button is pressed to assign that filter as the target response at the particular location (Fig. 7.2). When all adjustments are complete, the user clicks the “finished” button to return to the FDTD toolbox GUI. Once the target response GUI completes, the updated responses are fed into the remainder of the CSA algorithm and correction filters are generated as discussed in Chapter 5.

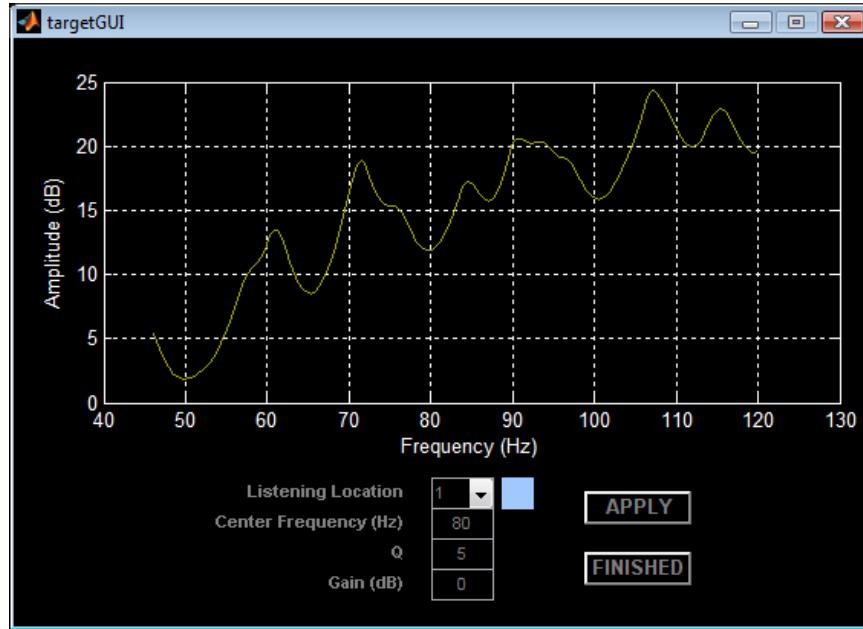


Fig. 7.1 Location-specific CSA equalization GUI (showing room average response)

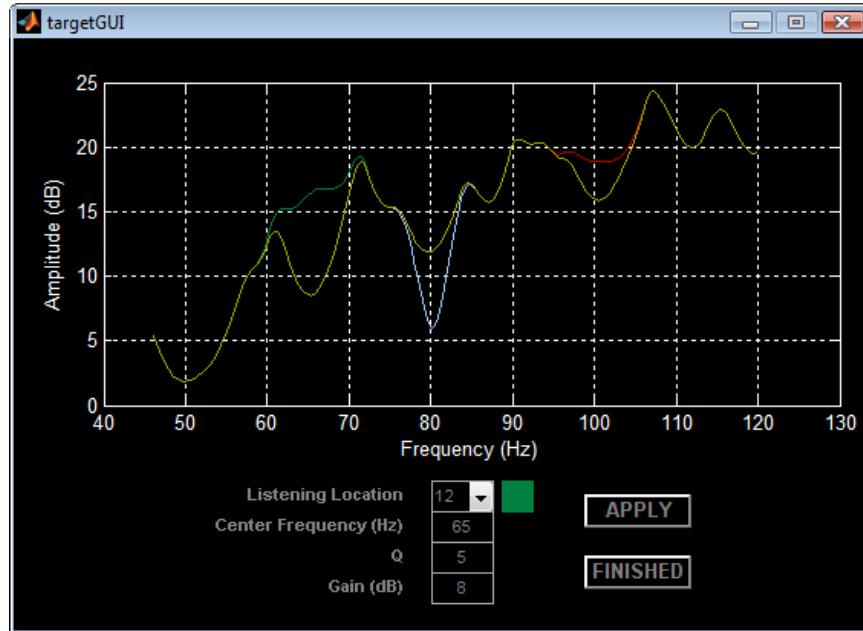


Fig. 7.2 Example of three adjusted CSA target frequency responses

The multiple-band CSA correction procedure lends only the higher frequency band ($\sim 40 - 120$ Hz) towards individual correction, as the lower band (below 40 Hz) utilizes only the omnidirectional components, reducing the number of target points by a factor of four. Fortunately, most home theater systems are in rooms where only the lowest room-mode falls below 40 Hz, resulting in room pressurization over the band with low spatial variance. If an omnidirectional subwoofer CSA were utilized, however, individual response correction can be applied over the entire subwoofer band due to the single CSA correction band in this case, but is limited by the efficient reproduction range of the subwoofers and also the size of the listening area. A smaller target listening area does not allow accurate spatial sampling at lower frequencies (due to wavelength) and therefore cannot accurately correct in the lower subwoofer band (see Chapter 5.2.2 for a complete discussion on this topic).

7.1.2 Response manipulation examples

A 5 m x 4 m x 3 m rectangular room was modeled with a hybrid subwoofer at each room floor corner. A sixteen-point listener grid was arranged in the room with its center at (2.4 m, 1.8 m, 1.8 m) and the frequency response at each location determined with a maximum length sequence (MLS). Uncorrected, the system exhibits a spatial variance of 5.58 dB (Fig. 7.3a) while after CSA correction with all target frequency responses identical, spatial variance is reduced by 96.3% to 0.21 dB (Fig. 7.3b). In this form, the system achieves an approximately equal response for all listeners, unlike the uncorrected system where the responses are clearly position-dependent.

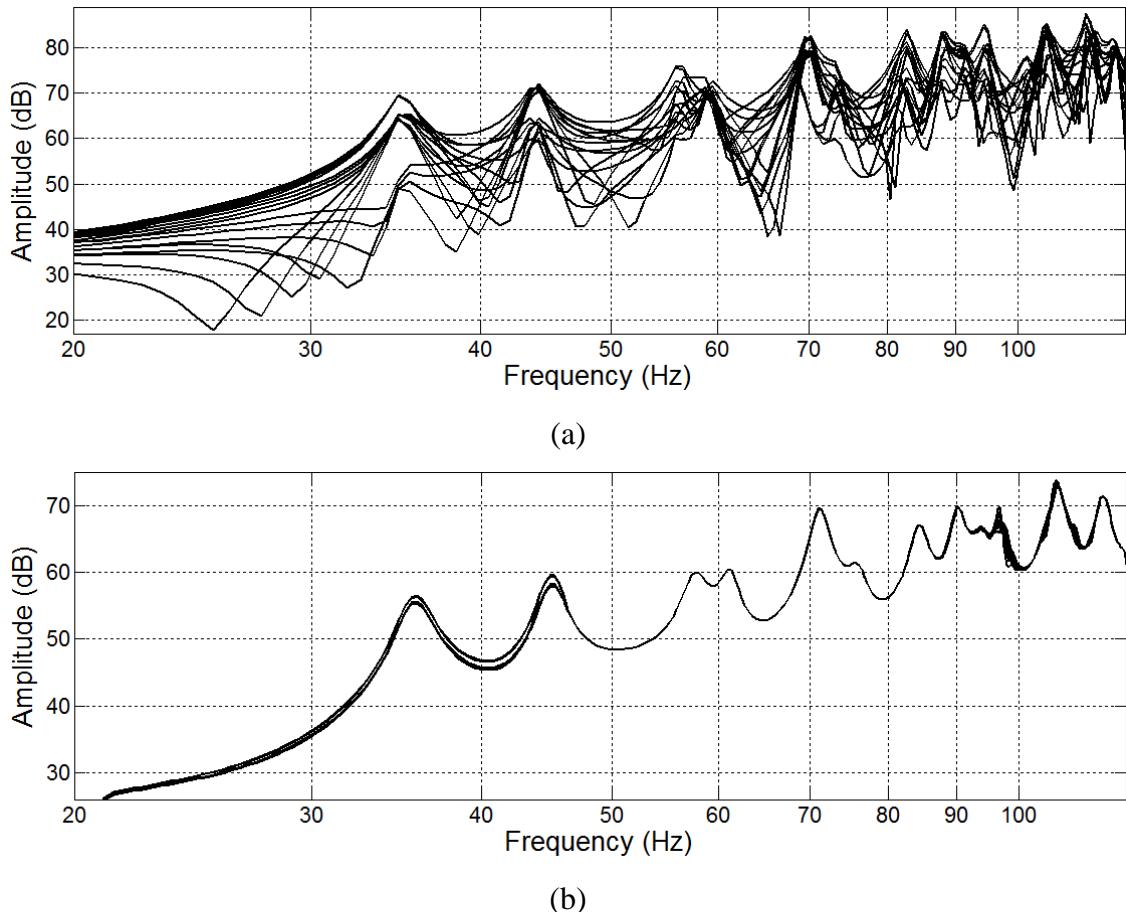


Fig. 7.3 System frequency responses (a) uncorrected and (b) after standard CSA correction

An identical configuration as that used for frequency-response correction (Fig. 7.3) is utilized to demonstrate position-specific equalization. Ideally, altering the response at one location should have no effect on other points. As an example, the resonance at 71 Hz can be attenuated for a single location (Fig. 7.4a). The single-location manipulation has no noticeable effect on any other point in the listening area, although system output has reduced by approximately 2 dB due to the added requirements on the correction filters. Another two listening locations can be adjusted to give a 9 dB boost centered at 80 Hz and a 9 dB attenuation centered at 59 Hz (Fig. 7.4b). All non-manipulated points maintain the room-average target frequency response (identical to that in Fig. 7.3b).

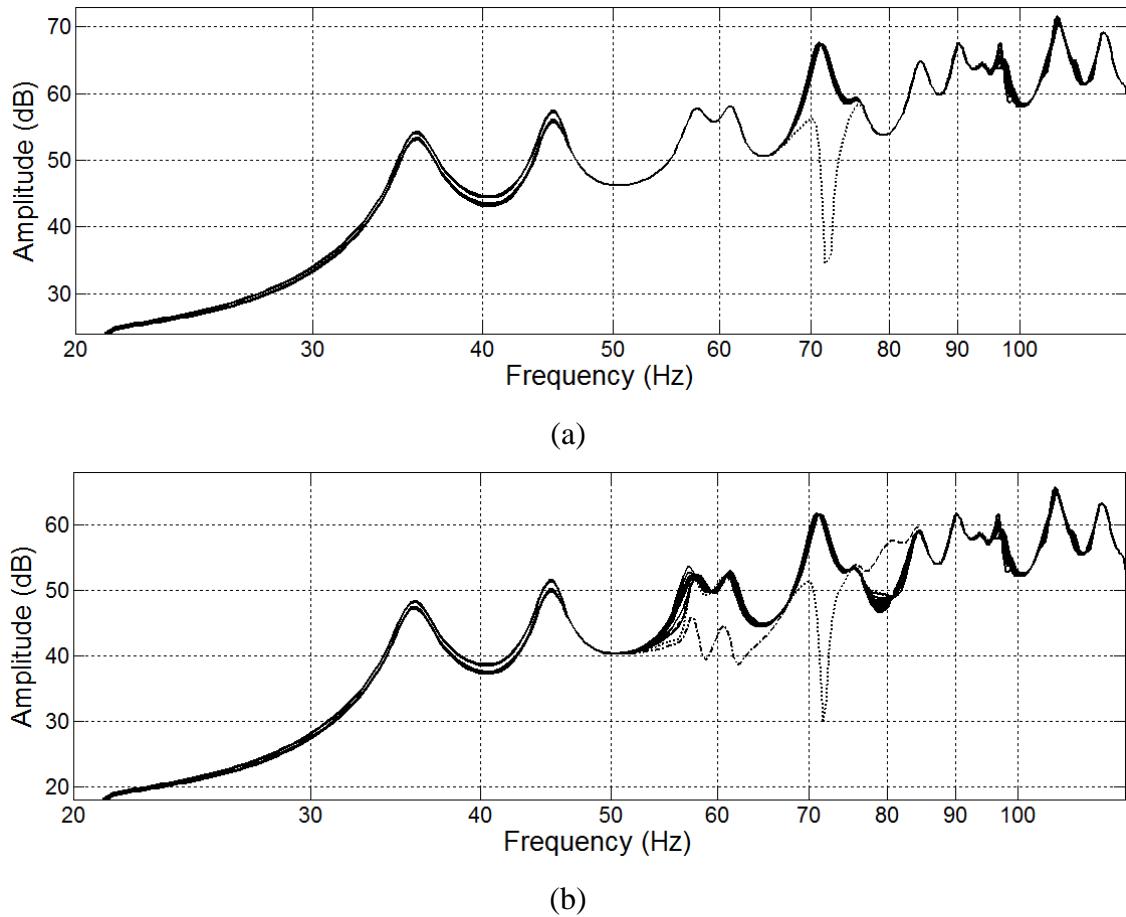


Fig. 7.4 Sixteen responses over a listening area with (a) one location manipulation: 71 Hz attenuation (dotted line) and (b) three location manipulations: 71 Hz attenuation, 80 Hz boost and 59 Hz attenuation (dotted, dashed and dash-dot lines, respectively)

The additional adjustments show the boost and attenuation at the target locations with only slight (< 1 dB) changes at other points over those frequency bands. The added requirements on the correction filters cause the system output to drop by 4 dB. It can be deduced that for every target point attempting to deviate from the measured room average target response, there will be approximately 2 dB loss in output. This issue calls for a sound reproduction system to have sufficient headroom to compensate for this reduction.

This individualized frequency control method can be visualized using spatio-pressure plots of the listening area at a given frequency. As an example, a similar configuration to that in the previous simulation can be used where a single listener requests the removal of the resonance at 71 Hz. A visualization of the sound pressure level at 71 Hz over the listening area is plotted for examination of how the procedure for a single listener affects surrounding listeners (Figure 7.5).

The example with the top right corner listener deciding to significantly attenuate 71 Hz highlights a number of important aspects of this procedure. First, as previously suggested, the output level reduces from around 93 dB with the uniformly-corrected CSA to 91 dB for the manipulated system. This supports the -2 dB per individual manipulation observed in Fig. 7.4. Additionally, the plots in Fig. 7.5 indicate that the different response in the top right corner does not affect the adjacent listeners (except with a reduced overall level which can be automatically adjusted when the manipulation request is received).

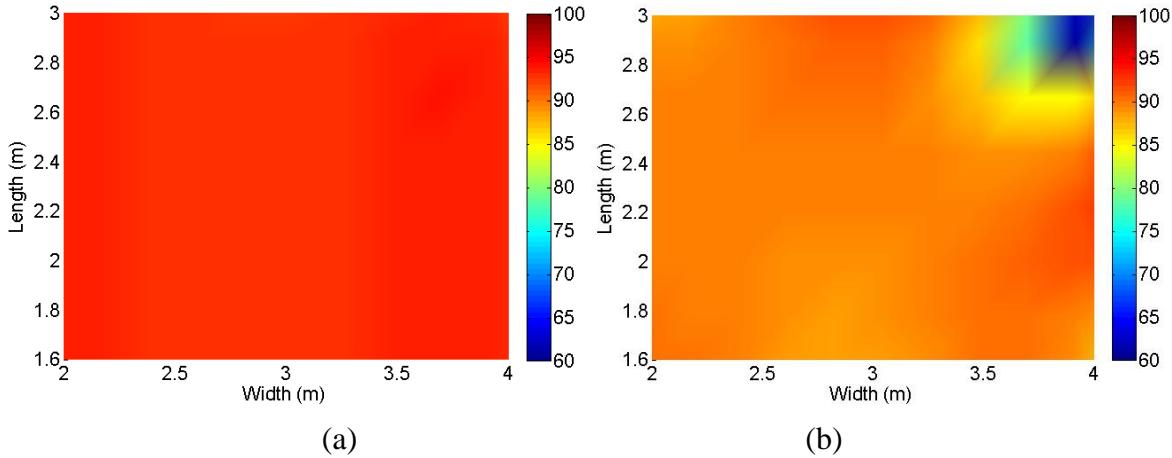


Figure 7.5 Spatio-pressure plots over the listening area (SPL in dB) at 71 Hz for (a) uniformly corrected CSA and (b) individually manipulated CSA correction (manipulation at the top right listener location)

In this case there are sixteen listening locations with the nearest locations to the point in question located at (3.4 m, 3.0 m), (3.4 m, 2.5 m) and (4.0 m, 2.5 m). The area influenced by the manipulated response ends just before the neighboring listening points, thus demonstrating listening response independence. Wavelength ultimately sets the effective lower limit of this procedure (coincident with the dipolar unit cutoff and overall listening area), but this test shows that even at 71 Hz, where the wavelength is 4.83 m, individualized correction can be applied to an area smaller than 0.3 m^2 which is significantly less than the manipulated frequency's wavelength.

These examples highlight the potential of CSAs not only for achieving an even response over the listening area, but also for individual control where multiple independent low-frequency responses can simultaneously exist in a relatively small area.

7.1.3 Discussion

A modified approach to CSA spatiotemporal low-frequency room control has been introduced, demonstrating how the frequency response at each listening location can be individually controlled while maintaining nearly complete independence from other locations. The effective frequency range of correction is limited to the higher operating band of the CSA, which is defined by measurement point spacing and the Schroeder frequency. In the case studied in this work, this amounted to a controllable range from 40 – 120 Hz over sixteen listening locations. Omnidirectional subwoofer CSAs do not exhibit this limitation (but will be limited by the subwoofers' reproduction capabilities and target point spacing).

System efficiency must be kept in mind, as a 2 dB drop in output sound pressure level was observed for each additional target response modification. A sound reproduction system with adequate headroom, however, should be able to handle this effect without many problems, but a filter optimization routine may prove beneficial in practical implementations as a required 3 dB output boost corresponds to a doubling of power. Although the presented examples show a 2 dB output loss for each response manipulation, further work is necessary to better understand this phenomenon.

As these examples exist entirely within the virtual domain, future work also necessitates practical experimentation. Even so, simulations give strong indication that this method can be practically realized, which could prove extremely useful for sound reproduction systems for

home theater, cinema, nightclub, live sound and even art installation applications where both objective and subjective criteria can be met, giving all listeners and/or performers a pleasing listening experience. Of course, as room size increases the independently-controllable areas will likely increase in size unless the quantity of degrees of freedom is increased.

Specifically, this technique may be useful in live sound where an even distribution of low-frequency energy over the audience area is required while simultaneously minimizing the energy on stage (discussed in Section 7.2). Along similar lines, a system of this sort could be applied in a nightclub where there can be designated “quiet” zones where low-frequency energy is minimized to give patrons a break from overly-strong bass, which can be fatiguing.

Beyond the zone-based control, individual seat-to-seat control is envisioned whereby listeners (in home theater, cinema or formal concert settings) access the correction system via a wireless device, such as a smart phone, and enter their seat number which allows them to adjust their localized response. This system, of course, requires a large number of degrees of freedom to allow for precise manipulation over a large audience. This requirement could be eased by taking additional measurements over the listening area and performing pseudo-inverse filtering within the CSA algorithm (as discussed in Chapter 5). As this is not the primary concern of this research, however, this investigation has been designated as future work.

7.2 Live sound

A high-quality sound system should provide consistent coverage over the entire audience area while at the same time keeping the sound pressure level (SPL) on stage to a minimum in order to provide musicians, technicians and production staff with a safe working environment. These requirements have predominantly been met with recent advancements in line array technology, where horizontal coverage patterns of 90° or 120° are readily achievable, minimizing sound wraparound to the stage. Coverage patterns in the subwoofer operating range (generally below 100 Hz), however, are more difficult to control using simple “one size fits all” system configurations.

Historically, industry-standard subwoofers have operated as approximately omnidirectional sources. In recent years, a handful of companies have introduced cardioid or supercardioid pattern subwoofers which can help limit low-frequency energy on stage, depending on the system configuration [92, 93]. In addition to this, technicians have long used the technique of rotating every other or every third subwoofer in vertical stacks 180° to achieve an approximate cardioid radiation pattern [94].

Conventional subwoofer systems suffer from a number of constraints, which differ from venue to venue including placement issues, rigging capabilities, sightlines, truck space and, of course, budget. These drawbacks can severely diminish a system’s capability to meet the low-frequency coverage and rejection criteria (a detailed discussion on these issues is presented in a conference paper reproduced in Appendix D). With this in mind, it is proposed that an adaptation of CSA technology can circumvent these practical issues and achieve venue-specific coverage patterns that benefit both the audience and stage areas.

7.2.1 Chameleon subwoofer array implementation

The CSA correction procedure is ideally suited for multi-component subwoofers, as described in Chapter 5; however, CSA correction can operate on any system given that each degree of freedom can be independently controlled. Cardioid subwoofers are increasingly common in the live sound industry, containing two independently-controlled drive-units.

The two distinct drive signals are principally generated within the system's control unit(s) which take the full-range input from the mixing desk and split the signal into relevant operating bands for each system component. CSA control is applied by inserting an extra DSP unit in between the system processor and the power amplifiers. This supplementary unit applies the control filters to the drive signals to create the target coverage pattern. Since the setup measurements are taken in-line with the system processor, the procedure will not be adversely affected by any processing unit manipulations.

As with small-room CSA applications, the live sound CSA procedure is limited by the number of degrees of freedom within the system and also the spacing of the target points. A subwoofer system driven by a four-mix amplifier rack on both sides of the stage gives a total of sixteen available degrees of freedom due to independent processing/amplification for front and rear drive signals (two degrees of freedom per subwoofer).

Frequencies with half-wavelengths shorter than the mean spacing between target points will not result in a uniform coverage pattern across the audience/stage area; rather pockets of control will arise caused by the wide spacing of target points. Small room CSA correction procedures generally recommend a more conservative one-quarter wavelength maximum spacing, but half-wavelength spacing is sufficient for the less detailed live sound CSA procedure.

A 50 m x 30 m virtual outdoor venue was set up containing an eight-unit cardioid subwoofer system with left/right placement. Simulations were conducted in two-dimensions to provide run-time efficiency. The system coverage pattern was simulated at 40 and 90 Hz for the unprocessed system (Fig. 7.6). An equally-spaced unprocessed system was also simulated for comparison purposes (Fig. 7.7).

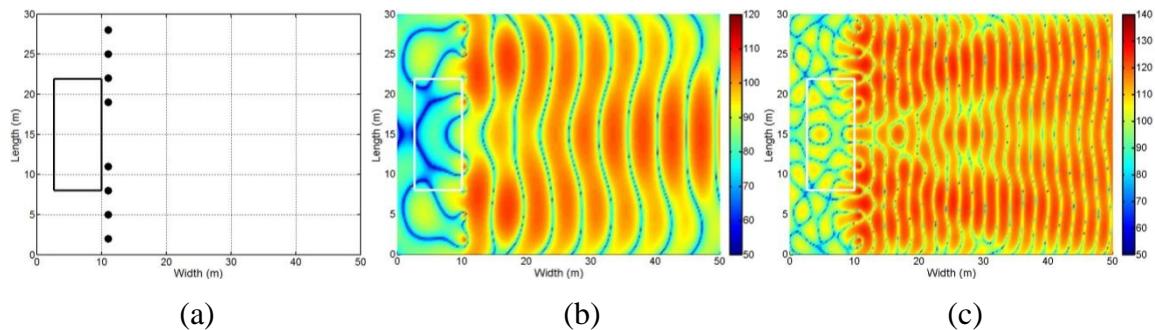


Fig. 7.6 Unprocessed left/right eight-unit cardioid subwoofer system with system layout (a) and simulated coverage patterns at (b) 40 Hz and (c) 90 Hz

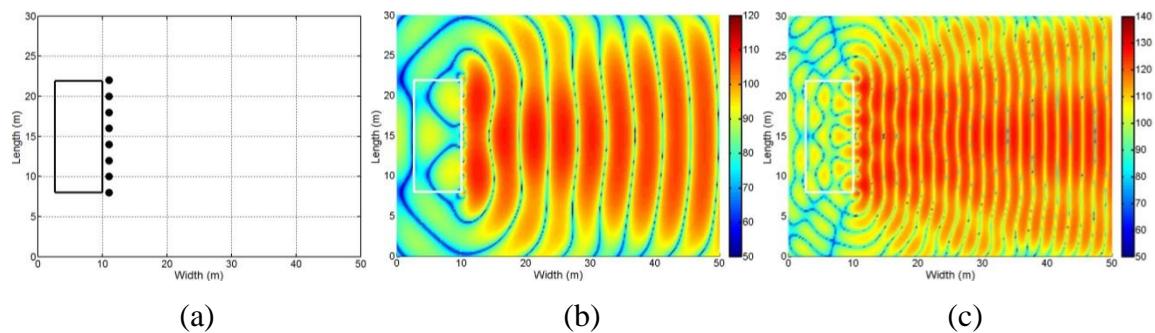


Fig. 7.7 Unprocessed central eight-unit cardioid subwoofer system with system layout (a) and simulated coverage patterns at (b) 40 Hz and (c) 90 Hz

The unprocessed systems give reasonable audience coverage patterns, although the left/right configuration results in spatial nulls at points within the audience and both systems deliver considerably less low-frequency energy towards the sides of the audience area. Both systems give approximately 10 – 15 dB stage rejection, as compared to SPL in the audience. At loud rock concerts this could amount to an SPL of over 110 dB at some areas of the stage, which is an unsafe daily working condition for musicians and production staff.

A CSA can be utilized in an attempt to minimize SPL on stage to ensure a safe working environment. Identical systems as in Figs. 7.6 and 7.7 were utilized for the CSA methodologies with target points arranged in a grid pattern on the stage. All target responses were set to the measured average response, but with approximately 40 dB attenuation. In addition, individual propagation delay was introduced at each target response, based on measured delay time to each target point. Audience coverage is not considered in this approach, although system output is set to match the uncorrected system, so similar audience coverage patterns should endure. The eight-cardioid unit systems grant sixteen degrees of freedom (pseudo-inverse filtering to allow for more measurement points is considered in the following section). Both configurations were tested at 40 and 90 Hz (Figs. 7.8 & 7.9).

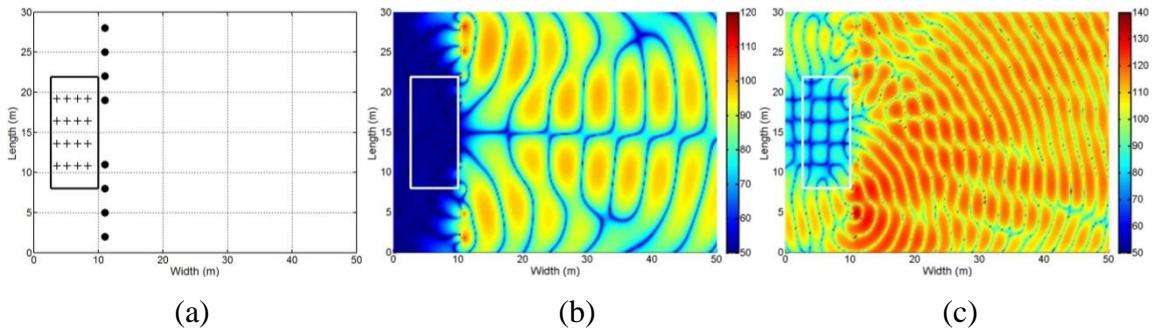


Fig. 7.8 CSA stage controlled left/right eight-unit cardioid subwoofer system with system layout (a) and simulated coverage patterns at (b) 40 Hz and (c) 90 Hz (+ = target point)

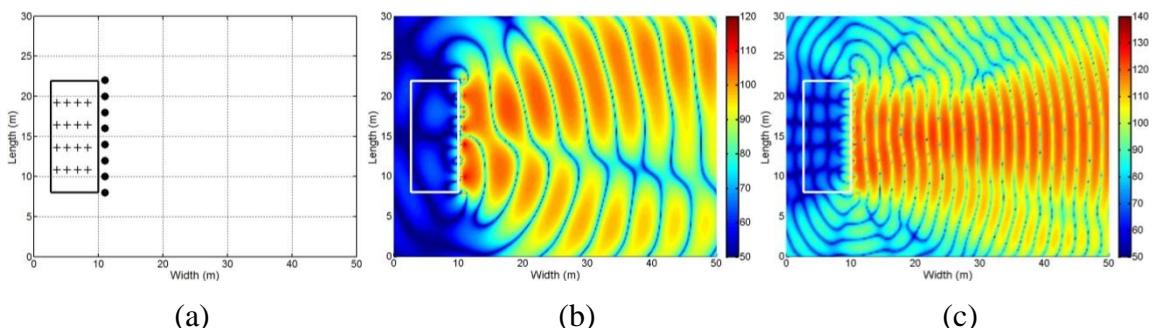


Fig. 7.9 CSA stage controlled central eight-unit cardioid subwoofer system with system layout (a) and simulated coverage patterns at (b) 40 Hz and (c) 90 Hz (+ = target point)

The stage-only CSA control results in significantly greater stage rejection (up to 50 dB in some cases), but also highlights certain problems. Since the audience area is not considered in the control procedure, there are clear deviations from the unprocessed coverage patterns, most noticeably with large central nulls at 40 Hz. The 90 Hz trials show less stage rejection than at 40 Hz due to the target point spacing. The mean spacing is 3 m which dictates that accurate, uniform control over the target area only exists below approximately 60 Hz. At 90 Hz, therefore, there are pockets of control with the outlying areas not benefiting from the CSA procedure.

To evade the problems present with stage-only CSA control, the sixteen target points can be divided between the stage and audience area. The stage target responses remain unchanged from the previous examples, while the audience points target the measured average response. This technique should provide more even audience coverage while maintaining stage rejection (although to a lesser extent due to fewer stage target points). Again, control is limited by point spacing to around 60 Hz (Figs. 7.10 & 7.11).

The stage + audience controlled CSA shows notable improvement over the stage-only control. The audience target points ensure that the coverage pattern is evenly distributed throughout the audience area while maintaining minimal pressure on stage. The downstage edge experiences relatively high SPL due to the target points' upstage placement. Again, at 90 Hz there are pockets of control as opposed to uniform control which is due to the wide target spacing, limiting uniform control to below 60 Hz. The central configuration, however, does result in even wide-audience coverage at 90 Hz, largely due to the naturally even coverage of the configuration.

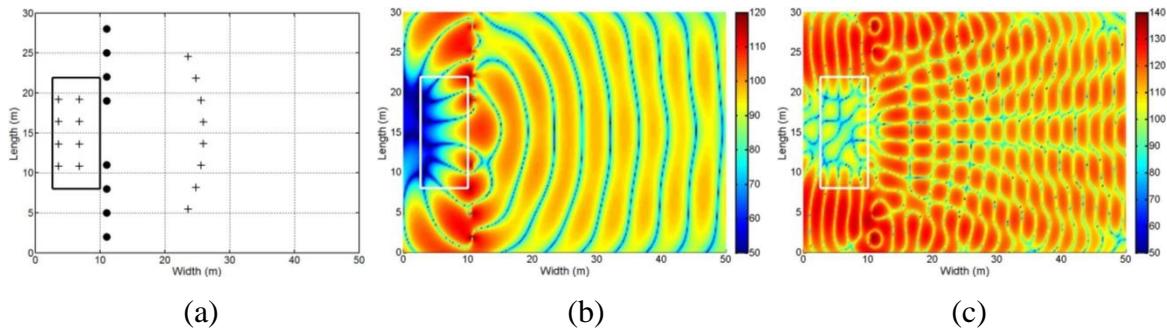


Fig. 7.10 CSA stage + audience controlled left/right eight-unit cardioid subwoofer system with system layout (a) and simulated coverage at (b) 40 Hz and (c) 90 Hz (+ = target point)

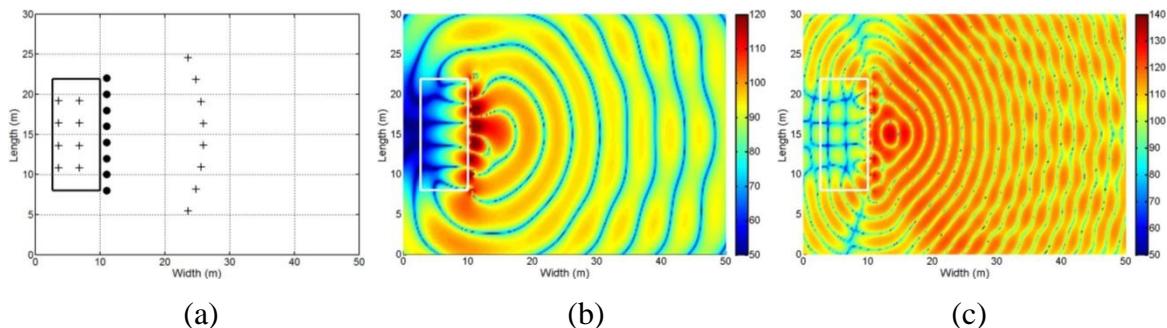


Fig. 7.11 CSA stage + audience controlled central eight-unit cardioid subwoofer system with system layout (a) and simulated coverage at (b) 40 Hz and (c) 90 Hz (+ = target point)

7.2.2 Pseudo-inverse filtering for live sound CSAs

A major drawback to the live sound CSA implementation in the previous section is the pockets of correction when there is inadequate spatial sampling of the stage and/or audience area (see Fig. 7.10a). This limits correction to only the lower end of the subwoofer band (below 60 Hz in the presented example).

A similar problem was encountered in Chapter 5 because there were too few source components, resulting in an inadequate set of target measurement points over the listening area. This problem was solved by utilizing pseudo-inverse filtering which allows for an unequal number of source components and target points (see Section 5.5 for details). With this

capability in place, listening areas can be properly spatially sampled, allowing for precise correction even while using a CSA with few degrees of freedom.

This approach can be applied to live sound CSAs to eliminate pockets of correction and give a smoother response across the stage and audience. The same configurations from Figs. 7.6 – 7.11 were simulated with the addition of pseudo-inverse matrix processing capabilities. Both the left/right and the central configurations were simulated using one-hundred target points split evenly between the stage and audience (Figs. 7.12 & 7.13).

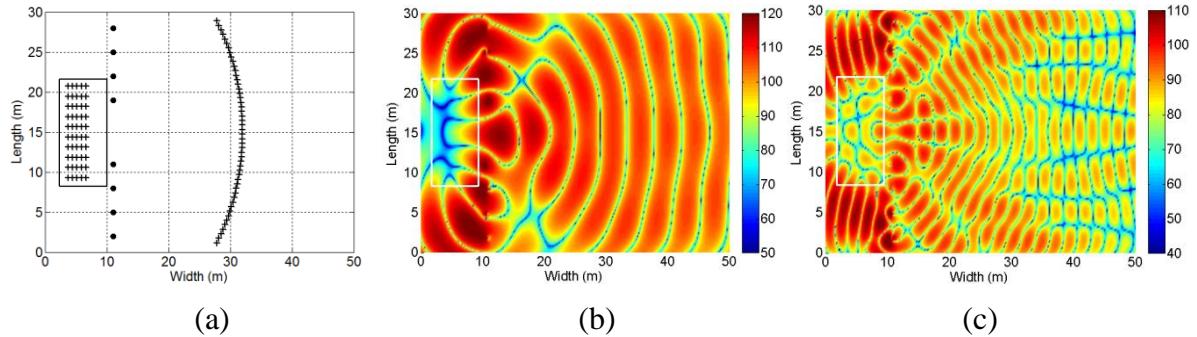


Fig. 7.12 CSA stage + audience controlled left/right eight-unit cardioid subwoofer system (using 100 target points using pseudo-inverse processing) with system layout (a) and simulated coverage at (b) 40 Hz and (c) 90 Hz (+ = target point)

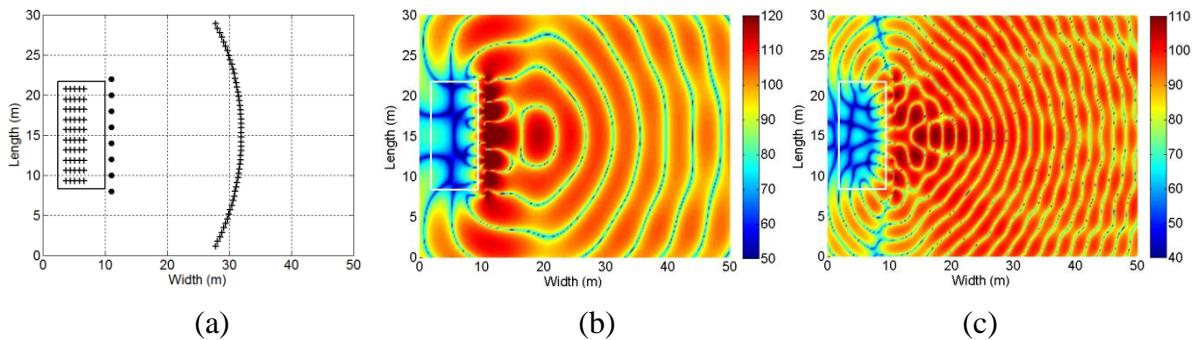


Fig. 7.13 CSA stage + audience controlled central eight-unit cardioid subwoofer system (using 100 target points using pseudo-inverse processing) with system layout (a) and simulated coverage at (b) 40 Hz and (c) 90 Hz (+ = target point)

The coverage patterns, although similar to those presented in Section 7.2.1, show specific improvements. First, the 40 Hz audience coverage pattern for the left/right configuration (Fig. 7.12b) is markedly wider than when only sixteen target points are utilized (Fig. 7.10b). Also, the 90 Hz audience coverage pattern (Fig. 7.12c) no longer shows pockets of correction since the target points are now spaced significantly closer to one another. Stage rejection in this case is not as high as hoped and the audience converge pattern still suffers from nulls due to the left/right subwoofer placement. These nulls, however, are largely absent in front of the arc of target points (from 10 – 30 m along the x-axis). This is a good indication that the system is still appropriately correcting for the target points. The area behind these points does not receive many correction benefits, which should be kept in mind if implementing a system of this sort.

The central subwoofer configuration (Fig. 7.13a) shows further improvement to both the stage and audience coverage. Due to the additional target points on stage, the area is better sampled than with the sixteen target point scenario (Fig. 7.11) which results in greater and more consistent stage rejection (Figs. 7.13b & 7.13c). The audience coverage pattern is largely

unchanged from the non-pseudo-inverse simulations due to the naturally even coverage of the subwoofer layout.

Clearly, pseudo-inverse filtering is as valuable in live sound scenarios as in small-room situations. Concerning the practicality of this method, a compromise is likely. Time generally will not permit one-hundred measurements in the audience and on stage. A systems engineer must determine how many measurements can reasonably be taken. If the subwoofer system crossover point is low enough (i.e. below 85, which is not uncommon) target points can have greater spacing, necessitating less time for calibration.

7.2.3 Discussion

A live sound application of chameleon subwoofer array (CSA) low-frequency control has been presented as a new technique to limit the amount of low-frequency energy on stage while creating a more uniform pressure distribution over a wide audience area. While conventional control systems provide up to around 20 dB attenuation on stage, largely due to the cardioid radiation pattern of the subwoofer units, CSA-controlled systems should be capable of up to 50 dB of rejection while maintaining uniform audience coverage.

Live sound CSA control operates within the existing framework of industry-standard systems, utilizing the independent front and rear drive-unit processing/amplification capabilities to allow for two degrees of freedom for each cardioid subwoofer. This method can be inserted into a conventional system in between the loudspeaker processor and the power amplifiers, minimizing the need for expensive new hardware.

A drawback to the CSA system is the target point spacing restrictions. Wider spacing corresponds to a lower control frequency limit. Above this limit control exists in pockets, but will not be uniform. Tighter target point spacing raises the frequency limit, but results in a narrower control area unless additional target points are added, which would require additional degrees of freedom within the subwoofer system. Alternatively, a pseudo-inverse filter calculation method can be implemented which improves correction area homogeneity, but may require significant time for the calibration measurements.

Future work required in this area includes simulating indoor environments to judge the effectiveness in smaller concert halls and other venues. As the system can easily fit into existing sound systems, real-world testing can theoretically be performed to confirm the effectiveness of the CSA live sound control procedure. Development of a more efficient measurement system can be targeted in future work whereby a microphone should be able to automatically scan areas of interest without the need for manual placement by a system technician.

Overall, CSA control can potentially bestow a robust solution to the difficult problem of providing equal low-frequency audience coverage while restricting SPL onstage. It differs from existing methodologies in that all units within the CSA operate synchronously to give the desired coverage pattern while other systems use independently operating subwoofers, requiring time-consuming manual fine tuning from venue to venue.

7.3 Conventional home theater

Section 7.2 highlights an alternative implementation of CSA technology for large-scale sound reinforcement. The proposal exploits the higher degrees of freedom normally available within a live sound subwoofer system, therefore only necessitating DSP to generate and apply the CSA

correction filters. This innate flexibility is likely to be attractive for adoption within the live sound industry as it requires minimal additional hardware.

A similar line of reasoning applies to home theater systems. Surround sound systems have become increasingly common in homes throughout the world since the mid-1970s. The most popular surround format is 5.1, consisting of front left, front center, front right, left surround and right surround loudspeakers with a single subwoofer for low-frequency effects (LFE) [91]. This setup is formalized by the ITU-R recommendation BS 775-1 (Fig. 7.14).

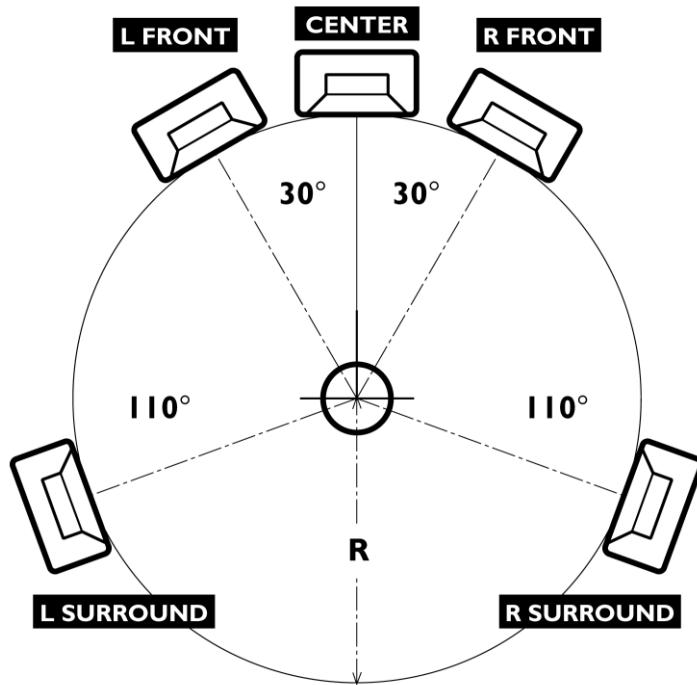


Fig. 7.14 5.1 surround sound configuration (ITU-R BS 775-1) [91]

Assuming all non-subwoofer loudspeakers in a 5.1 system can reproduce sound efficiently to low frequencies (though not necessarily extreme bass where there is only room pressurization) then the system permits six degrees of freedom. The non-subwoofers can be treated like the dipole pairs in hybrid subwoofers (as discussed in Chapter 5), where their correction range has a lower limit based on the loudspeakers' cutoff frequency. An omnidirectional subwoofer, of course, is ideally efficient over the entire low-frequency range.

While this class of CSA implementation cannot be expected to achieve as precise results as subwoofer arrays due to the frequency limitations of the non-subwoofer loudspeakers, it can eliminate the strongest areas of spatial variance, thus providing more consistent low-frequency coverage across the listening area. The number of available degrees of freedom is system dependant as surround systems sometimes utilize smaller surround and center loudspeakers which cannot operate efficiently below roughly 80 Hz.

7.3.1 Example scenario

A virtual 5.1 surround sound home theater configuration was arranged in the FDTD toolbox using a 5 m x 4 m x 3 m room. The omnidirectional subwoofer was placed at the front left room corner. Following the established practice, a 9-point listening grid was arranged to encompass the primary listening location (as defined in Fig. 7.14), shown in Fig. 7.15.

The non-subwoofer loudspeakers were modeled to exhibit a -6 dB/octave roll-off below 200 Hz to approximate their limited low-frequency reproduction capabilities. The uncorrected system gives a spatial variance of 4.50 dB over the listening area (Fig. 7.16).

CSA correction was implemented with the system crossover point set to 62 Hz so that nothing below that point is fed to the non-subwoofer loudspeakers. Again, the target response was set to the room average. Figure 7.17 shows the CSA-corrected system results.

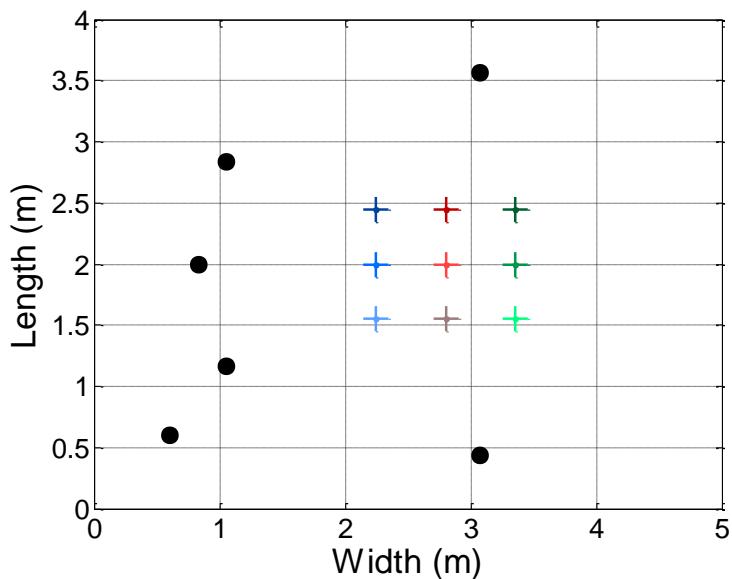


Fig. 7.15 Simulated 5.1 surround sound configuration

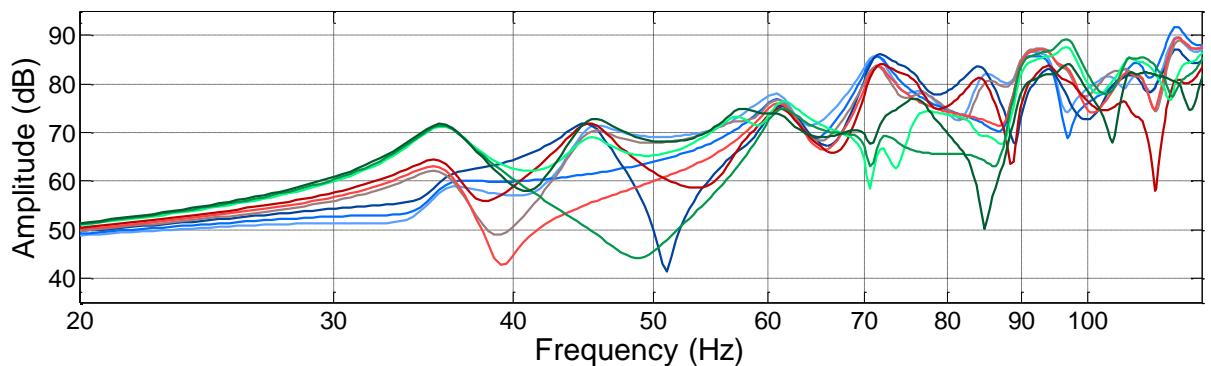


Fig. 7.16 Uncorrected 5.1 surround sound system frequency responses

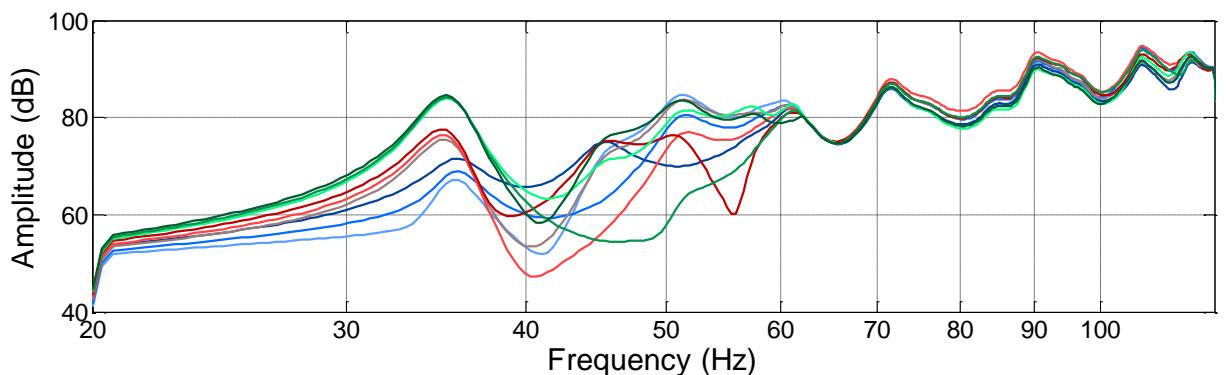


Fig. 7.17 CSA-corrected 5.1 surround sound system frequency responses

The post-correction results highlight a number of characteristics of this form of CSA. First, there is no correction benefit in the low-band of the system (below 62 Hz). Only the subwoofer is active in this range and contributes just a single degree of freedom which cannot provide control at multiple points, as discussed in Chapter 6. Above the crossover frequency, the system behaves as desired with six degrees-of-freedom. Overall, spatial variance has been reduced to 2.64 dB; a reduction of 41.3%.

7.3.2 Discussion

As with the live sound example, CSA correction is practical within standard surround sound home theater systems. The CSA operates by treating all loudspeakers as part of the array, each contributing one degree of freedom. The correction band must therefore be split into two ranges: one lower range for the subwoofer only and an upper range exploiting the full complement of loudspeakers. This strategy avoids overloading the non-subwoofer loudspeakers at very low-frequencies.

Simulated results confirm accurate control over the upper band, which in the example case ranges from 62 – 120 Hz. The lower subwoofer band receives no correction benefits since there is only a single available degree of freedom. This may not prove too great compromise as film soundtracks tend to rely on the subwoofer channel only for transitory effects such as explosions. The bulk of the aural information during the program is likely to exist above this range, where the CSA can provide adequate control. The surround sound accuracy should be maintained with CSA processing since most directional cues are located in the non-subwoofer band. This of course requires subjective verification if the system were to be practically implemented. In cases where a surround sound system is utilized for high quality music reproduction, higher accuracy in the lower subwoofer range is more critical and could benefit from a hybrid subwoofer. This suggestion is an appropriate topic for future investigations.

7.4 Chapter summary

Although the primary motivation for development of the CSA low-frequency room response control technology is to achieve minimal sound pressure variance over a wide-area in a small room, there are additional applications. The three discussed in this chapter focus on system flexibility in terms of target correction responses as well as embedding the CSA algorithm within existing industry-standard hardware (both in small-room home theater applications and large-scale live sound scenarios).

Individualized low-frequency response manipulation is achieved with a CSA containing sufficient degrees of freedom to address each listener location. Every target point can theoretically possess its own unique target correction response, thus allowing individuals to adjust the response to their liking in real-time. This idea was demonstrated using an FDTD simulation with three points out of a sixteen point listener grid deviating from the default target response. The individualized modifications were achieved without adversely affecting other locations. The only downside shown was an approximate 2 dB loss in output level for each response modification when deviating from the default. This should not prove a problem with systems containing sufficient power amplification and high-quality drive-units.

Since it is unlikely that the average consumer would invest in additional hardware (either more omnidirectional or hybrid subwoofers) to meet the degrees of freedom requirements of CSAs, a modified CSA algorithm can be applied to a standard surround sound home theater system. Although not all loudspeakers may reproduce efficiently at the lowest frequencies, it can be

assumed that they operate acceptably at least partially into the subwoofer band. This is highlighted with the simulation of a standard 5.1 surround sound system. The correction band is split into two ranges: the lower handled exclusively by the subwoofer and the higher handled by all system components. In the example case the crossover point is 62 Hz. Since the low-band possesses a single degree of freedom, no correction benefits are realized below 62 Hz; however the high-band contains six degrees of freedom which is shown to provide precise control, minimizing spatial variance.

Since the very-low-frequency band in surround sound film soundtracks is primarily utilized for effects such as explosions, the spatial variance in this range may not be as critical. Additionally, the human ear is less sensitive to amplitude deviations in this frequency range [95, 96], therefore the variance may not be as audible. Consequently, this CSA application is considered attractive especially as DSP can readily be loaded into home theater processing units.

Lastly, CSAs in live sound can be useful in solving common sound reinforcement issues. This application embraces aspects of the individualized target response approach and the surround sound implementation. As with the surround sound approach, the CSA control exists entirely within DSP, without the need for special hardware. Since modern-day live sound subwoofer systems commonly utilize cardioid subwoofers, containing two degrees of freedom each, a large-scale sound system provides considerable degrees of freedom.

If the target points are split between the stage and the audience area, there are two regions of correction. In the audience the target responses are set to the average measured response to achieve equal coverage over the entire audience. On the other hand, the target responses on stage are set to approximately 40 dB below the measured average response over the audience area to provide a safe working environment for all musicians and stage personnel.

Simulations have shown that this approach is feasible although care must be taken when selecting the target listening points. If these points are too widely spaced, the corrected system provides only pockets of correction instead of a consistent response over the area. This necessitates a greater number of closely-spaced measurement points to meet the system goals over a large area or, alternatively, the inclusion of a pseudo-inverse filtering routine where more measurements are taken than there are degrees of freedom and an optimization routine determines the best-fit solution.

These three extended CSA applications have been shown through simulation to be feasible and to provide greater system flexibility. The latter two systems can be applied economically within existing DSP-enabled hardware.

Overall CSA technology has been shown to have applications in a range of situations. The system is naturally flexible and is applicable to a number of different scenarios, each with varying low-frequency coverage goals.

8 Conclusions

The core objective of this research was to develop a robust method for accurate and efficient spatiotemporal control of low-frequency room-modes in small-sized listening rooms. Three interdependent subject areas were covered to meet the stated goals: acoustical modeling (Chapter 3), low-frequency room-mode correction (Chapters 4 & 5) and virtual bass (Chapter 6). All three topics are rooted in the general area of low-frequency acoustics (Chapter 2). Although achieving minimal-variance of low-frequency sound fields in small-room applications was the principal goal of the research, this spawned extended applications embracing user-specific frequency response manipulation and the implementation of the derived correction algorithm into both existing hardware structures of home-theater surround sound systems and large-scale live sound reinforcement systems (Chapter 7).

8.1 Critical assessment

The core original contributions of this research span three major components: the FDTD simulation toolbox, chameleon subwoofer arrays and virtual bass systems. Each aspect of these achievements has been disseminated in journals (Appendix C) and presented at conferences (Appendix D).

8.1.1 FDTD simulation toolbox

The motivation for the FDTD simulation toolbox was that publically or commercially available acoustical modeling software did not exist in appropriate form which allowed for sufficient low-frequency resolution and flexibility in terms of system configuration, including the application of correction/equalization procedures. Chapter 4 and Appendices A and B comprehensively demonstrate the capabilities of the toolbox which greatly expedited this research.

The FDTD method results in appropriately accurate low-frequency data with efficient runtime. The simplified boundary conditions were chosen as opposed to more accurate locally-conformal boundaries since absolute accuracy was not necessary, rather the toolbox was meant to generate results to indicate the strengths and weaknesses of various correction procedures. The simulation results were validated rigorously through comparison to measurements, theoretical modal values and previously published work highlighting different simulation software. These comparisons confirmed proper toolbox functionality, allowing broader research to move forward utilizing the toolbox as a virtual testing ground.

The simulation features, including non-rectangular topologies, auralization, sound propagation animation/visualization, variable subwoofer polar patterns, user-definable source signals and the various optimization routines allow simulations of nearly any scenario. The extensive data analysis features, which provide detailed investigations of both spectral and transient responses at numerous measurement points, have been shown to be a powerful research aid.

At the time of writing, the toolbox has been made available as a freely-downloadable open-source project (www.adamjhill.com/fdtd) with the anticipation of receiving external feedback and suggestions for improvement. The toolbox therefore remains a work in progress with room for improvement and expansion (as discussed in Section 8.2.1). The toolbox has proven an integral research tool, facilitating timely completion of various modeling requirements while presenting an accessible avenue for experimentation, which for example has led to the emergence of the chameleon subwoofer array concept.

8.1.2 Chameleon subwoofer arrays

Chameleon subwoofer arrays (CSA), initially prototyped within the FDTD simulation toolbox, were developed to provide maximal sound field control within the listening area of a small-sized room. The specific goal at the outset was to minimize spatial variance among listeners in order to maintain objectively similar frequency responses at all listening locations. Complementary to this requirement was the desire to maintain accurate time-domain performance of the pressure response over a wide listening area as typified for example through tone-burst testing. In addition, it was observed that CSAs also allow for listener-specific frequency response manipulations which address user-subjectivity. Listeners initially receive the same objective response, but can adjust this response to meet their specific listening preferences. As this feature was tangential to the core research, future work is required to refine the method (as detailed in Section 8.2.2).

Inspiration for CSAs was gathered from available research concentrating on spatial variance minimization. The idea of exploiting multiple independently-controllable system components (i.e. degrees of freedom) to accurately control a sound field is not in itself new. For example Welti [19] utilized numerous omnidirectional subwoofers and a least-mean-squares (LMS) brute-force optimization routine to determine a set of equalization filters that achieves a maximally flat room response.

However, brute-force optimization presents a problem for a system allowing for user-defined target responses (as with CSAs). The LMS procedure was noted by Welti to often take minutes to complete, which does not lend itself well to real-time adjustments. Additionally, the system operated using omnidirectional subwoofers, necessitating many individual units placed throughout a room which, as has been discussed in this thesis, can be impractical. Nevertheless, it was Welti and Toole who proposed four subwoofers as the upper limit for practical systems [2, 17] where their work has greatly motivated CSA development.

Backman [5] has demonstrated the potential of frequency-dependant subwoofer directivity. Backman observed that at very low-frequencies (at or below the lowest room-mode), directional subwoofers such as dipoles become inefficient while omnidirectional units can retain room pressurization. Backman inserts an all-pass filter between the dipole pair which provides the required phase-shift to convert the unit into an omnidirectional radiator below a defined frequency. This allows subwoofers to minimally excite key room-modes in the discrete modal band (using a directional pattern) while adequately pressurizing the room below the modal band with an omnidirectional pattern.

This necessitates the determination of what variety of polar pattern to implement for CSAs. The answer has its origins in work conducted by Howe and Hawksford on an ambisonics-style method for sound field equalization [73]. The method consisted of an ambisonics loudspeaker setup where a listener was surrounded by sources that operated in combination to generate the required sound field (using spherical harmonics). This idea was adapted and incorporated within the *hybrid subwoofer*. These subwoofers may contain six drive-units which contribute to four independently-controllable spherical harmonics (or degrees of freedom using the CSA terminology): one omnidirectional (using all six drive-units) and three dipoles. Using this methodology, the Welti-Toole proposal can be implemented with four subwoofers granting a total of sixteen, rather than four, degrees of freedom while eliminating the requirement of listeners to be surrounded by loudspeakers (as with the Howe-Hawksford method).

The CSA direct calculation procedure, while allowing for near-real-time frequency response adjustments, requires a series of constraints/bounds to ensure the calculated solution is practical; a point strongly emphasized by Welti [19]. These constraints include limiting the dipole component bandwidth to the upper subwoofer band (as determined by the room topology and the measurement point layout), monitoring the correction coefficients' amplitudes to avoid drive-unit over-excursion and windowing each correction filter impulse response to suppress persistent ringing.

Simulation has confirmed this concept capable of suppressing spatial variance over wide-areas. Naturally, the greater number of degrees of freedom within the system construct, the more spatiotemporal response control is available. Appropriate spatial sampling was shown to be critical within these simulations, as under-sampling the target area results in degraded performance. It was shown that pseudo-inverse filtering allows for additional measurements and largely solves this issue by ensuring proper spatial sampling while not requiring a large number of subwoofers.

A prototype hybrid subwoofer was constructed inspired by work conducted in the simulation stage of this research. This prototype was tested in a standard listening environment exploiting pseudo-inverse filtering to allow the listening area to be finely sampled. Experimental and simulated results proved to be in close agreement, indicating that the four degree of freedom hybrid subwoofer can deliver over 50% spatial variance reduction. Additionally, omnidirectional subwoofer-based CSAs were tested and were shown to exhibit consistent results.

Drawbacks to the CSA methodology have been revealed. Calibration measurements can be tedious because a large number of points are required for large listening areas. In home theater applications, this may call for around twenty-five measurements which consumers may not have the patience to perform. Additionally, CSAs show some problematic layout sensitivity. Concerning subwoofer placement, CSA simulations show that poor-performing passive layouts (largely due to close subwoofer placement) result in mean output level drops between the uncorrected and corrected systems. It cannot be assumed that a consumer-grade system would have the capabilities to compensate for this loss. Post-calibration subwoofer movement can also be problematic, especially if a unit is moved towards a room corner or the listening area. Movement parallel to the listening area shows less sensitivity.

Overall, CSAs combine a number of parallel research tracks concerning low-frequency room-mode correction. The hybrid subwoofer packs four source components into one unit, granting greater correction flexibility. The direct calculation approach gives users the ability to adjust target response(s) in near-real-time, a capability not available in many existing proposals. However, CSA signal processing can be applied to existing hardware such as surround sound systems in home theater or large-scale subwoofer systems for live sound applications. This should prove more appealing to consumers as it requires less investment in new equipment. Regardless of the chosen CSA implementation, the technique has closely met the stated goals at the outset and has highlighted required future work to improve system performance and practicality.

8.1.3 Virtual bass systems

Virtual bass was explored as a novel means to augment room-mode correction. The concept focuses on the fact that CSA correction (and indeed other methods) commonly exhibit difficulty in correction over a handful of narrow frequency bands. This problem reduces the system's

output efficiency as well as potentially causing drive-unit over-excursion. Virtual bass correction can remove these problematic bands from the physically reproduced signal and substitute them with synthesized virtual bass for subjective reinforcement.

Commencing the exploration into virtual bass it was noted that the two predominant synthesis mechanisms, nonlinear devices (NLS) and phase vocoders (PV), perform poorly for certain varieties of signals. To counteract these inherent drawbacks, a hybrid virtual bass synthesis approach was developed which dynamically monitors the transient content of the input signal and correspondingly weights the two mechanisms running in parallel. This approach was subjectively evaluated which indicated that the hybrid synthesis operates consistently over all tested genres of music (unlike the NLD and PV systems). It must be noted that the hybrid system does not necessarily outperform the NLD and PV systems in terms of sound quality ratings, but it gives more consistent sound quality ratings for all input signals, unlike the other methods.

The hybrid synthesis method was incorporated into a virtual bass room-mode correction technique which was tested as a standalone correction procedure by targeting five of the strongest room-modes in a listening room. The tests were designed to rate the sound quality at two adjacent seats as well as the observed low-frequency seat-to-seat variance. This test was repeated after virtual bass processing and showed the procedure to significantly reduce perceived spatial variance amongst listeners without compromising system fidelity.

An overuse of virtual bass can produce synthetic sounding results, but because this procedure is only supplemental to existing correction techniques, it largely mitigates this problem. This system was demonstrated in context of CSA applications where a combination of the measurement matrix condition number and correction coefficient amplitude was used to identify frequency bands for virtual bass replacement. Simulation has confirmed improved efficiency with CSAs due to the relaxed reproduction requirements; however future work is still necessary to fully-integrate virtual bass processing within the CSA prototype.

8.2 Future work

This thesis focuses on the development of CSAs primarily through well-described theory and detailed simulations. The prototype CSA used for practical experimentation indicates that this methodology is indeed feasible and performs as predicted by the simulations. To move this concept (and/or the individual developments stemming from the FDTD simulation toolbox and the hybrid virtual bass system) towards a broader (possibly commercial) appeal, further research is necessary.

8.2.1 FDTD simulation toolbox

The simulation toolbox is fully-functional and reliable software although it could benefit from additional work focusing on the following:

- Improve runtime efficiency by streamlining the MATLAB code. Tasks include removing unnecessary nested *for* loops and eliminating repeated code.
- Frequency-dependent absorption implementation using IIR filters for each wall element. Surface element memory permits frequency-dependent relationships, but requires care to avoid efficiency issues due to the additional processing requirements.

- Expand boundary conditions to the locally-conformal method. This would increase the simulation's accuracy by permitting rounded boundaries, resulting in less scattering errors. As with the absorption modifications, this requires consideration of computational complexity to keep runtimes to a minimum.
- If wideband simulations are required, the FDTD modeling can be joined with a geometrical technique to handle high-frequency simulations. This would require a careful implementation to properly interface with the non-rectangular topology configuration procedure, which is designed specifically for use with FDTD modeling.

8.2.2 Chameleon subwoofer arrays

Future work on chameleon subwoofers should concentrate on improvements to the hybrid subwoofer, but must also address individualized response manipulation, applications within existing hardware and real-time implementations. The generalized focus areas are as follows:

- Hybrid subwoofer development
 - Alter the source component configuration so that the omnidirectional component utilizes all drive-units (as with the simulations) and each dipole component operates using two opposing drivers.
 - Investigate source component cross-talk, which has been postulated as a significant factor in correction error. A solution might include a cross-talk suppression matrix applied to the output of the DSP routine.
 - Examine the frequency response of the subwoofer. Does it efficiently reproduce down to 20 Hz or is a form of low-frequency extension necessary? Could porting the subwoofer work within the CSA procedure?
- Chameleon subwoofer array signal processing routine
 - Conduct a thorough investigation concerning the window shape/size applied to the correction filter impulse responses. Should the window be fixed or defined by the correction filters? In some cases can windowing be omitted?
 - Expand the CSA routine to automatically identify problematic sets of filters and (potentially) apply an optimization routine to determine the best practical solution (while not significantly increasing processing time).
 - Improve the individualized correction procedure by minimizing output efficiency losses (possibly through an optimization routine) and testing within the prototype system.
- Chameleon subwoofer array applications
 - Test within various home theater surround sound systems (experiment with both small and large surround speakers).
 - Test within a live sound reinforcement system, both for audience and stage correction benefits.
 - Refine the calibration procedure and eventually convert the routine to a real-time system and test.

- Perform listening evaluations of CSA correction systems blindly compared to other correction techniques and various uncorrected systems.
- Low-frequency directionality
 - Investigate whether low-frequency directionality within the context of a small listening space is possible using a CSA.
 - Perform simulations tracking the velocity vectors of sound waves and the received signals at each listening location to generate a clearer picture of how directional information is received.
 - Conduct outdoor and indoor blind listening tests for subwoofer directionality and determine (and confirm with the reviewed literature) how well humans determine the direction of a low-frequency sound in anechoic/non-anechoic environments. Do high-frequency cues override low-frequency cues?

8.2.3 Virtual bass systems

The virtual bass correction method can be expanded in the following areas:

- Fully-integrate the procedure within CSA processing and perform subjective testing.
- Refine the hybrid synthesis approach (primarily the transient content detector) to generate the best-possible virtual bass effect.

8.3 Closing remarks

The research presented in this thesis critically analyzes established low-frequency modeling and correction methods as well as common virtual bass synthesis procedures in order to develop a novel small-room low-frequency room-mode correction system. The *FDTD simulation toolbox* allows for extremely flexible acoustical modeling involving various sound reproduction system including DSP-based correction methods.

Virtual prototyping using the toolbox gave rise to *chameleon subwoofer arrays* (CSA) which have proven capable of precise low-frequency spatiotemporal response control over wide listening areas. These arrays (along with the accompanying DSP) additionally can provide individualized frequency response manipulation for each listener and are implementable within existing hardware such as home theater surround sound and live sound reinforcement systems. A supplemental virtual bass room-mode correction system was developed to handle narrow frequency bands that prove difficult to correct. The addition of this procedure has shown to improve output efficiency and therefore increases the practicality of CSAs.

A comprehensive set of guidelines for CSA implementation has been laid out with relevant simulation and experimental results included to support the various claims. The concept can benefit from future work concentrating on fine-tuning the hybrid subwoofer (especially concerning drive-unit cross-talk) and CSA implementations within existing sound reproduction systems and/or other multiple drive-unit subwoofers. Perusing this work will permit CSA technology to be more accessible within the mainstream and could possibly lead to various commercially-viable products.

References

- [1] Rossing, T.D.; F.R. Moore; P.A. Wheeler. *The Science of Sound, 3rd Edition*. Addison Wesley, New York. 2002.
- [2] Toole, F.E. *Sound Reproduction: Loudspeakers and Rooms*. Focal Press, New York. 2008.
- [3] Walker, R. "Low-frequency room responses: Part 1 – Background and qualitative considerations." BBC Research Department Report. RD 1992/8. 1992.
- [4] Sabine, W.C. *Collected Papers on Acoustics*. Harvard University Press, Cambridge, MA. 1923.
- [5] Backman, J. "Low-frequency polar pattern control for improved in-room response." 115th Convention of the Audio Engineering Society, paper 5867. October, 2003.
- [6] Benjamin, E.; B. Gannon. "Effect of room acoustics on subwoofer performance and level setting." 109th Convention of the Audio Engineering Society, paper 5232. September, 2000.
- [7] Walker, R. "Low-frequency room responses: Part 2 – Calculation methods and experimental results." BBC Research Department Report. RD 1992/9. 1992.
- [8] Bradley, J.S.; G.A. Soulodre; Scott Norcross. "Factors influencing the perception of bass." 133rd ASA/Noise-Con 97 Meeting. June, 1997.
- [9] Kuttruff, H. "Sound fields in small rooms." 15th International AES Conference, paper 15-002. October, 1998.
- [10] Olson, H.F. *Music, Physics and Engineering, 2nd Edition*. Dover Publications, Inc. New York. 1967.
- [11] Fletcher, H. "On the dynamics of the cochlea." Journal of the Acoustical Society of America, vol. 23, no. 6, pp. 637-645. November, 1951.
- [12] Greenwood, D.D. "Critical bandwidth and the frequency coordinates of the basilar membrane." Journal of the Acoustical Society of America, vol. 33, no. 10, pp. 1344-1356. October, 1961.
- [13] Schroeder, M.R. "Die statistischen oarameter der frequenzkurve von großen raumen." Acustica, vol. 4, pp. 594-600. 1954.
English translation: M.R. Schroeder. "Statistical parameters of the frequency response curves of large rooms." Journal of the Audio Engineering Society, vol. 35, no. 5, pp. 299-306. May, 1987.
- [14] Schroeder, M.R. "The 'Schroeder frequency' revisited." Journal of the Acoustical Society of America, vol. 99, no. 5, pp. 3240-3241. May, 1996.
- [15] Geddes, E.R. "Small room acoustics in the statistical region." 15th International Conference of the Audio Engineering Society. October, 1998.
- [16] Schroeder, M.R; K.H. Kuttruff. "On frequency response curves in rooms. Comparison of experimental, theoretical, and Monte Carlo results for the average frequency spacing

- between maxima." Journal of the Acoustical Society of America, vol. 34, no. 1, pp. 76-80. January, 1962.
- [17] Welti, T. "How many subwoofers are enough." 112th Convention of the Audio Engineering Society, paper 5602. May, 2002.
- [18] Celestinos, A.; S. Birkedal Nielsen. "Optimizing placement and equalization of multiple low frequency loudspeakers in rooms." 119th Convention of the Audio Engineering Society, paper 6545. October, 2005.
- [19] Welti, T.; A. Devantier. "Low-frequency optimization using multiple subwoofers." Journal of the Audio Engineering Society, vol. 54, no. 5, pp. 347-364. May, 2006.
- [20] Waterhouse, R.V. "Output of a sound source in a reverberation chamber and other reflecting environments." Journal of the Acoustical Society of America, vol. 30, no. 1, pp. 4-13. January, 1958.
- [21] Vanderkooy, J. "Applications of the acoustic centre." 122nd Convention of the Audio Engineering Society, paper 7102. May, 2007.
- [22] Vanderkooy, J. "The low-frequency acoustic centre: Measurement, theory and application." 128th Convention of the Audio Engineering Society, paper 7992. May, 2010.
- [23] Hubbard, H.H; K.P. Shepherd. "The Helmholtz resonance behaviour of single and multiple rooms." NASA Contractor Report #178173. September, 1986.
- [24] Putland, G.R. "Acoustical properties of air versus temperature and pressure." Journal of the Audio Engineering Society, vol. 42, no. 11, pp. 927-933. November, 1994.
- [25] Kemp, J.A. "Theoretical and experimental study of wave propagation in brass musical instruments, Chapter 7 – Improvements to pulse reflectometry." Ph.D. Thesis, University of Edinburgh, UK. 2002.
- [26] Kemp, J.A. "Improvements to bore profile measurement in acoustic pulse reflectometry." Proc. Institute of Acoustics, vol. 30, pt. 2. 2008.
- [27] Hill, A.J. "Voice coil temperature in loudspeaker performance: Causes, effects and prevention techniques." M.Sc. Dissertation, University of Edinburgh, UK, 2008.
- [28] Corrington, M.S. "Correlation of transient measurements on loudspeakers with listening tests." Journal of the Audio Engineering Society, vol. 3, no. 1, pp. 35-39. January, 1955.
- [29] Kaminsky, V.J. "The response of loudspeaker to tone bursts." Journal of the Audio Engineering Society, vol. 13, no. 2, pp. 119-123. April, 1965.
- [30] Linkwitz, S. "Narrow band impulse testing of acoustical systems." 60th Convention of the Audio Engineering Society, paper 1342. May, 1978.
- [31] Linkwitz, S. "Shaped tone-burst testing." Journal of the Audio Engineering Society, vol. 28, no. 4, pp. 250-258. April, 1980.
- [32] Linkwitz, S. "Investigation of sound quality differences between monopolar and dipolar woofers in small rooms." 105th Convention of the Audio Engineering Society, paper 4786. September, 1998.

- [33] Rindel, J.H. "The use of computer modeling in room acoustics." *Journal of Vibroengineering*, vol. 3, no. 4, pp. 219-224. 2000.
- [34] Gibbs, B.M.; D.K. Jones. "A simple image method for calculating the distribution of sound pressure levels within an enclosure." *Acustica*, vol. 26, pp. 24-32. 1972.
- [35] Allen, J.B.; D.A. Berkley. "Image method for efficiently simulating small-room acoustics." *Journal of the Acoustical Society of America*, vol. 64, no. 4, pp. 943-950. April, 1979.
- [36] Krokstad, A.; S. Strom; S. Sorsdal. "Calculating the acoustical room response by the use of a ray tracing technique." *Journal of Sound and Vibration*, vol. 8, pp. 118-125. 1968.
- [37] Bansal, M.; S. Feistel; W. Ahnert. "Large scale FEM analysis of a studio room." 120th Convention of the Audio Engineering Society, paper 6736. May, 2006.
- [38] Celestinos, A.; S. Birkedal Nielsen. "Multi-source low frequency room simulation using finite difference time domain approximations." 117th Convention of the Audio Engineering Society, paper 6264. October, 2004.
- [39] Henwood, D.J. "The boundary-element method and horn design." *Journal of the Audio Engineering Society*, vol. 41, no. 6, pp. 485-496. June, 1993.
- [40] Karjalainen, M.; A. Jdrvinen; L. Savioja; J. Backman; V. Ikonen; A. Suutala; S. Pohjolainen. "Comparison of numerical simulation models and measured low-frequency behavior of a loudspeaker." 104th Convention of the Audio Engineering Society, paper 4722. May, 1998.
- [41] Botteldooren, D. "Acoustical finite-difference time-domain simulation in quasi-cartesian grid." *Journal of the Acoustical Society of America*, vol. 95, pp. 2313-2319. May, 1994.
- [42] Olesen, S.K. "Low frequency room simulation using finite difference equations." 102nd Convention of the Audio Engineering Society, paper 4422. March, 1997.
- [43] Shannon, C.E. "Communication in the presence of noise." *Proceedings of the I.R.E.*, vol. 37, pp. 10-21. January, 1949.
- [44] Kowalczyk, K.; M. van Walstijn. "Modelling frequency-dependent boundaries as digital impedance filters in FDTD and K-DWM acoustics simulations." 124th Convention of the Audio Engineering Society, paper 7430. May, 2008.
- [45] Suzuki, H.; A. Omoto; K. Fujiwara. "Treatment of boundary conditions by finite time domain method." *Acoustical Science and Technology*, vol. 1, no. 8, pp. 16-26. 2007.
- [46] Tolan, J.G.; J.B. Schneider. "Locally conformal method for acoustic finite-difference time-domain modeling of rigid surfaces." 140th Meeting of the Acoustical Society of America. July, 2003.
- [47] Clio 8.5, Audiomatica. www.audiomatica.com/clio/overview
- [48] Walker, R. "Optimum dimension ratios for small rooms." 100th Convention of the Audio Engineering Society, paper 4191. May, 2002.
- [49] Milner, J.R.; R.J. Bernhard. "An investigation of the modal characteristics of nonrectangular reverberation rooms." *Journal of the Acoustical Society of America*, vol. 85, no. 2, pp. 772-779. February, 1989.

- [50] Hufker, B. "Fibonacci numbers, the 'golden mean,' and audio engineering." *dB*, pp. 41-43. June, 1979.
- [51] Huylebrouck, D.; P. Labarque. "More true applications of the golden number." *Nexus Network Journal*, vol. 4, no. 1, pp. 44-58. 2002.
- [52] Markowsky, G. "Misconceptions about the golden ratio." *The College of Mathematics Journal*, vol. 23, no. 1, pp. 2-19. January, 1992.
- [53] Fletcher, J.A. "The design of a modular sound absorber for very low frequencies." BBC Research Department, report number RD1992/10. 1992.
- [54] Reiley, E.; A. Grimani. "Room mode bass absorption through combined diaphragmatic & Helmholtz resonance techniques: 'The Springzorber.'" 114th Convention of the Audio Engineering Society, paper 5760. March, 2003.
- [55] Preis, D. "Phase distortion and phase equalization in audio signal processing - A tutorial review." 105th Convention of the Audio Engineering Society, paper 1849. October, 1981.
- [56] Genereux, R.P. "Adaptive loudspeaker systems: correcting for the acoustics environment." 8th International Conference of the Audio Engineering Society. May, 1990.
- [57] Kuriyama, J.; Y. Furukawa. "Adaptive loudspeaker system." *Journal of the Audio Engineering Society*, vol. 37, no. 11, pp. 919-926. November, 1989.
- [58] Pedersen, J.A.; K. Hermansen. "The distribution of the low frequency sound field and its relation to room equalization." 96th Convention of the Audio Engineering Society, paper 3852. February, 1994.
- [59] Elliot, S.J.; P.A. Nelson. "Multiple-point equalization in a room using adaptive digital filters." *Journal of the Audio Engineering Society*, vol. 37, no. 11, pp. 899-907. November, 1989.
- [60] Bharitkar, S. "Phase equalization for multi-channel loudspeaker-room responses." 117th Convention of the Audio Engineering Society, paper 6272. October, 2004.
- [61] Bharitkar, S.; C. Kyriakakis. "Comparison between time delay based and non-uniform phase based equalization for multi-channel loudspeaker-room responses." 119th Convention of the Audio Engineering Society, paper 6607. October, 2005.
- [62] Bharitkar, S.; C. Kyriakakis. "Objective function and automatic multi-position equalization and bass management filter selection." 119th Convention of the Audio Engineering Society, paper 6608. October, 2005.
- [63] Santillan, A.O. "Experimental low-frequency sound equalization in an extended region of an enclosure using adaptive filters." 104th Convention of the Audio Engineering Society, paper 4755. May, 1998.
- [64] Santillan, A.O. "Spatially extended sound equalization in rectangular rooms." *Journal of the Acoustical Society of America*, vol. 110, no. 4, pp. 1989-1997. October, 2001.
- [65] Elliot, S.J.; L.P. Bhatia; D.W. Deghan; A.H. Fu; M.S. Stewart; D.W. Wilson. "Practical implementation of low-frequency equalization using adaptive digital filters." *Journal of the Audio Engineering Society*, vol. 42, no. 12, pp. 988-998. December, 1994.

- [66] Olson, H.F. "Gradient Loudspeakers." *Journal of the Audio Engineering Society*, vol. 21, no. 2, pp. 86-93. March, 1973.
- [67] Ferekidis, L.; U. Kempe. "The beneficial coupling of cardioid low frequency sources to the acoustic of small rooms." 116th Convention of the Audio Engineering Society, paper 6110. May, 2004.
- [68] Darlington, P.; M.R. Avis. "Time/frequency response of a room with active acoustic absorption." 100th Convention of the Audio Engineering Society, paper 4192. May, 1996.
- [69] Celestinos, A.; S. Birkedal Nielsen. "Low frequency sound field enhancement system for rectangular rooms using multiple low frequency loudspeakers." 120th Convention of the Audio Engineering Society, paper 6688. May, 2006.
- [70] Vanderkooy, J. "Multi-source room equalization: Reducing room resonances." 123rd Convention of the Audio Engineering Society, paper 7262. October, 2007.
- [71] Makivirta, A.; P. Antsalo; M. Karjalainen; V. Valimaki. "Modal equalization of loudspeaker-room responses at low frequencies." *Journal of the Audio Engineering Society*, vol. 51, no. 5, pp. 324-434. May, 2003.
- [72] Pedersen, J.A. "Adjusting a loudspeaker to its acoustic environment – The ABS system." 115th Convention of the Audio Engineering Society, paper 5880. October, 2003.
- [73] Howe, R.M.; M.O.J. Hawksford. "Methods of local room equalization and their effect over the listening area." 91st Convention of the Audio Engineering Society, paper 3138. October, 1991.
- [74] Fletcher, H.; W.A. Munson. "Loudness, its definition, measurement and calculation." *Journal of the Acoustical Society of America*, vol. 5, no. 2, pp. 82-108. October, 1933.
- [75] Penrose, R. "A generalized inverse for matrices." *Proc. Cambridge Philosophical Society*, vol. 51, pp. 406-413. 1955.
- [76] Mark of the Unicorn, Inc. "MOTU PCI-424 User's Guide for Windows." Cambridge, MA, USA. 2002. cdn-data.motu.com/manuals/pci-audio/PCI-424_User_Guide_Win.pdf
- [77] Audiomatica, SRL. "Signal Conditioner SC-02 Technical Specifications." Florence, Italy. 2009. www.audiomatica.com/download/depliantpciusb.pdf
- [78] Schouten, J.F.; R.J. Ritsma; B. Lopes Cardozo. "Pitch of the residue." *Journal of the Acoustical Society of America*, vol. 34, no. 8, pp. 1418-1424. September, 1962.
- [79] Oo, N.; W.S. Gan. "Analytical and perceptual evaluation of nonlinear devices for virtual bass systems." 128th Convention of the Audio Engineering Society, paper 8108. May, 2010.
- [80] Larsen, E.; R.M. Aarts. "Reproducing low-pitched signals through small loudspeakers." *Journal of the Audio Engineering Society*, vol. 50, no. 3, pp. 147-164. March, 2002.
- [81] Daniel, B.T.; C. Martin. "The effect of the MaxxBass psychoacoustic bass enhancement system on loudspeaker design." 106th Convention of the Audio Engineering Society, paper 4892. May, 1999.

- [82] Bai, M.R.; W.C. Lin. "Sythesis and implementation of virtual bass system with a phase-vocoder approach." *Journal of the Audio Engineering Society*, vol. 54, no. 11, pp. 1077-1091. November, 2006.
- [83] Robel, A. "A new approach to transient processing in the phase vocoder." *Proc. 6th Int. Conference on Digital Audio Effects*. London, UK. September, 2003.
- [84] Nagel, F.; S. Disch; N. Rettelback. "A phase vocoder driven bandwidth extension method with novel transient handling for audio codecs." *127th Conference of the Audio Engineering Society*, paper 7711. May, 2009.
- [85] Laroche, J.; M. Dolson. "New phase-vocoder techniques for real-time pitch shifting, chorusing, harmonizing, and other exotic audio modifications." *Journal of the Audio Engineering Society*, vol. 47, no. 11, pp. 928-936, November, 1999.
- [86] Zolzer, U. *DAFX: Digital Audio Effects*, pp244-256. John Wiley & Sons, Ltd. Chichester, West Sussex, UK. 2002.
- [87] Brown, J.C. "Calculation of a constant Q spectral transform." *Journal of the Acoustical Society of America*, vol. 89, no. 1, pp. 425-434, January, 1991.
- [88] Mason, A.J. "The MUSHRA audio subjective test method." *BBC Research & Development White Paper*, WHP 038, September, 2002.
- [89] Aarts, R.M. "Band stop filter." International Patent Number WO 00/57673. September 28, 2000.
- [90] Tan, C.T.; B.C.J. Moore; N. Zacharov. "The effect of nonlinear distortion on the perceived quality of music and speech signals." *Journal of the Audio Engineering Society*, vol. 51, no. 11, pp. 1012-1031. November, 2003.
- [91] Lund, T. "Enhanced localization in 5.1 production." *109th Convention of the Audio Engineering Society*, paper 5243. September, 2000.
- [92] Meyer Sound Laboratories, Inc. Berkeley, California, USA. www.meyersound.com
- [93] Nexo SA. Plailly, France. www.nexo-sa.com
- [94] Rat Sound Systems, Inc. Oxnard, California, USA. www.ratsound.com
- [95] Zwicker, E.; U.T. Zwicker. "Audio engineering and psychoacoustics: Matching signals to the final receiver, the human auditory system." *Journal of the Audio Engineering Society*, vol. 39, no. 3, pp. 115-125. March, 1991.
- [96] Buck, M. "Perceptual attributes of supraliminal low frequency sound and infrasound." *66th Convention of the Audio Engineering Society*, paper 1663. May, 1980.
- [97] Pierce, A.D. *Acoustics – An introduction to its physical principles and applications*. Acoustical Society of America, Melville, New York, USA. 1989.

Appendix A – FDTD simulation equations and mask set

The first two sections of this appendix contain a full set of the two- and three-dimensional equations utilized within the FDTD simulation toolbox. Equations for both non-boundary and boundary conditions are included as well as any additional relevant formulas. The third section of this appendix presents a full set of masks for an example two-dimensional non-rectangular model in the toolbox.

Additionally, a CD-ROM is included, attached to the rear page of this thesis, which contains both the publically-available (CSA functions removed) and the comprehensive (all functions included) versions of the FDTD toolbox and the virtual bass synthesis software along with the relevant user-guides. The CD-ROM also contains a digital version of this thesis as well as high-resolution color versions of all papers listed in Appendices C and D.

A.1 Two-dimensional FDTD simulation equations

Variable definition:

u^x	=	x-dimension particle velocity (m/s)
u^y	=	y-dimension particle velocity (m/s)
$p_{x,y}$	=	sound pressure at point (x, y) (Pa)
t	=	time (s)
Δt	=	time step (s)
Δx	=	element spacing in the x-direction (m)
Δy	=	element spacing in the y-direction (m)
Z	=	characteristic wall impedance
α	=	absorption coefficient
R_x, R_y	=	particle velocity constants
c	=	speed of sound in air (m/s)
ρ	=	air density (kg/m^3)

A.1.1 Non-boundary particle velocity

$$u_{x+\frac{\Delta x}{2},y}^x \left(t + \frac{\Delta t}{2} \right) = u_{x+\frac{\Delta x}{2},y}^x \left(t - \frac{\Delta t}{2} \right) - \frac{\Delta t}{\rho \Delta x} [p_{x+\Delta x,y}(t) - p_{x,y}(t)] \quad (A.1)$$

$$u_{x,y+\frac{\Delta y}{2}}^y \left(t + \frac{\Delta t}{2} \right) = u_{x,y+\frac{\Delta y}{2}}^y \left(t - \frac{\Delta t}{2} \right) - \frac{\Delta t}{\rho \Delta y} [p_{x,y+\Delta y}(t) - p_{x,y}(t)] \quad (A.2)$$

A.1.2 All pressure values

$$\begin{aligned} p_{x,y}(t + \Delta t) = & p_{x,y}(t) - \frac{c^2 \rho \Delta t}{\Delta x} \left[u_{x+\frac{\Delta x}{2},y}^x \left(t + \frac{\Delta t}{2} \right) - u_{x-\frac{\Delta x}{2},y}^x \left(t + \frac{\Delta t}{2} \right) \right] \\ & - \frac{c^2 \rho \Delta t}{\Delta y} \left[u_{x,y+\frac{\Delta y}{2}}^y \left(t + \frac{\Delta t}{2} \right) - u_{x,y-\frac{\Delta y}{2}}^y \left(t + \frac{\Delta t}{2} \right) \right] \end{aligned} \quad (A.4)$$

A.1.3 Characteristic wall impedance & particle velocity constants

$$Z = \rho c \frac{1 + \sqrt{1 - \alpha}}{1 - \sqrt{1 - \alpha}} \quad (A.5)$$

$$R_x = \frac{\rho \Delta x}{0.5 \Delta t} \quad (A.6)$$

$$R_y = \frac{\rho \Delta y}{0.5 \Delta t} \quad (A.7)$$

A.1.4 Boundary particle velocity (x-dimension)

Positive direction boundary condition:

$$u_{x+\frac{\Delta x}{2},y}^x \left(t + \frac{\Delta t}{2} \right) = \frac{R_x - Z}{R_x + Z} u_{x+\frac{\Delta x}{2},y}^x \left(t - \frac{\Delta t}{2} \right) + \frac{2}{R_x + Z} p_{x-\Delta x,y}(t - \Delta t) \quad (A.9)$$

Negative direction boundary condition:

$$u_{x+\frac{\Delta x}{2},y}^x \left(t + \frac{\Delta t}{2} \right) = \frac{R_x - Z}{R_x + Z} u_{x+\frac{\Delta x}{2},y}^x \left(t - \frac{\Delta t}{2} \right) - \frac{2}{R_x + Z} p_{x-\Delta x,y}(t - \Delta t) \quad (A.10)$$

A.1.5 Boundary particle velocity (y-dimension)

Positive direction boundary condition:

$$u_{x,y+\frac{\Delta y}{2}}^y \left(t + \frac{\Delta t}{2} \right) = \frac{R_y - Z}{R_y + Z} u_{x,y+\frac{\Delta y}{2}}^y \left(t - \frac{\Delta t}{2} \right) + \frac{2}{R_y + Z} p_{x,y-\Delta y}(t - \Delta t) \quad (A.11)$$

Negative direction boundary condition:

$$u_{x,y+\frac{\Delta y}{2}}^y \left(t + \frac{\Delta t}{2} \right) = \frac{R_y - Z}{R_y + Z} u_{x,y+\frac{\Delta y}{2}}^y \left(t - \frac{\Delta t}{2} \right) - \frac{2}{R_y + Z} p_{x,y-\Delta y}(t - \Delta t) \quad (A.12)$$

A.2 Three-dimensional FDTD simulation equations

Variable definition:

u^x	=	x-dimension particle velocity (m/s)
u^y	=	y-dimension particle velocity (m/s)
u^z	=	z-dimension particle velocity (m/s)
$p_{x,y,z}$	=	sound pressure at point (x, y, z) (Pa)
t	=	time (s)
Δt	=	time step (s)
Δx	=	element spacing in the x-direction (m)
Δy	=	element spacing in the y-direction (m)
Δz	=	element spacing in the z-direction (m)
Z	=	characteristic wall impedance
α	=	absorption coefficient
R_x, R_y, R_z	=	particle velocity constants
c	=	speed of sound in air (m/s)
ρ	=	air density (kg/m^3)

A.2.1 Non-boundary particle velocity

$$u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) = u_{x+\frac{\Delta x}{2},y,z}^x \left(t - \frac{\Delta t}{2} \right) - \frac{\Delta t}{\rho \Delta x} [p_{x+\Delta x,y,z}(t) - p_{x,y,z}(t)] \quad (\text{A.13})$$

$$u_{x,y+\frac{\Delta y}{2},z}^y \left(t + \frac{\Delta t}{2} \right) = u_{x,y+\frac{\Delta y}{2},z}^y \left(t - \frac{\Delta t}{2} \right) - \frac{\Delta t}{\rho \Delta y} [p_{x,y+\Delta y,z}(t) - p_{x,y,z}(t)] \quad (\text{A.14})$$

$$u_{x,y,z+\frac{\Delta z}{2}}^z \left(t + \frac{\Delta t}{2} \right) = u_{x,y,z+\frac{\Delta z}{2}}^z \left(t - \frac{\Delta t}{2} \right) - \frac{\Delta t}{\rho \Delta z} [p_{x,y,z+\Delta z}(t) - p_{x,y,z}(t)] \quad (\text{A.15})$$

A.2.2 All pressure values

$$\begin{aligned} p_{x,y,z}(t + \Delta t) = & p_{x,y,z}(t) - \frac{c^2 \rho \Delta t}{\Delta x} \left[u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) - u_{x-\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) \right] \\ & - \frac{c^2 \rho \Delta t}{\Delta y} \left[u_{x,y+\frac{\Delta y}{2},z}^y \left(t + \frac{\Delta t}{2} \right) - u_{x,y-\frac{\Delta y}{2},z}^y \left(t + \frac{\Delta t}{2} \right) \right] \\ & - \frac{c^2 \rho \Delta t}{\Delta z} \left[u_{x,y,z+\frac{\Delta z}{2}}^z \left(t + \frac{\Delta t}{2} \right) - u_{x,y,z-\frac{\Delta z}{2}}^z \left(t + \frac{\Delta t}{2} \right) \right] \end{aligned} \quad (\text{A.16})$$

A.2.3 Characteristic wall impedance & particle velocity constants

$$Z = \rho c \frac{1 + \sqrt{1 - \alpha}}{1 - \sqrt{1 - \alpha}} \quad (\text{A.17})$$

$$R_x = \frac{\rho \Delta x}{0.5 \Delta t} \quad (\text{A.18})$$

$$R_y = \frac{\rho \Delta y}{0.5 \Delta t} \quad (A.19)$$

$$R_z = \frac{\rho \Delta z}{0.5 \Delta t} \quad (A.20)$$

A.2.4 Boundary particle velocity (x-dimension)

Positive direction boundary condition:

$$u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) = \frac{R_x - Z}{R_x + Z} u_{x+\frac{\Delta x}{2},y,z}^x \left(t - \frac{\Delta t}{2} \right) + \frac{2}{R_x + Z} p_{x-\Delta x,y,z}(t - \Delta t) \quad (A.21)$$

Negative direction boundary condition:

$$u_{x+\frac{\Delta x}{2},y,z}^x \left(t + \frac{\Delta t}{2} \right) = \frac{R_x - Z}{R_x + Z} u_{x+\frac{\Delta x}{2},y,z}^x \left(t - \frac{\Delta t}{2} \right) - \frac{2}{R_x + Z} p_{x-\Delta x,y,z}(t - \Delta t) \quad (A.22)$$

A.2.5 Boundary particle velocity (y-dimension)

Positive direction boundary condition:

$$u_{x,y+\frac{\Delta y}{2},z}^y \left(t + \frac{\Delta t}{2} \right) = \frac{R_y - Z}{R_y + Z} u_{x,y+\frac{\Delta y}{2},z}^y \left(t - \frac{\Delta t}{2} \right) + \frac{2}{R_y + Z} p_{x,y-\Delta y,z}(t - \Delta t) \quad (A.23)$$

Negative direction boundary condition:

$$u_{x,y+\frac{\Delta y}{2},z}^y \left(t + \frac{\Delta t}{2} \right) = \frac{R_y - Z}{R_y + Z} u_{x,y+\frac{\Delta y}{2},z}^y \left(t - \frac{\Delta t}{2} \right) - \frac{2}{R_y + Z} p_{x,y-\Delta y,z}(t - \Delta t) \quad (A.24)$$

A.2.6 Boundary particle velocity (z-dimension)

Positive direction boundary condition:

$$u_{x,y,z+\frac{\Delta z}{2}}^z \left(t + \frac{\Delta t}{2} \right) = \frac{R_z - Z}{R_z + Z} u_{x,y,z+\frac{\Delta z}{2}}^z \left(t - \frac{\Delta t}{2} \right) + \frac{2}{R_z + Z} p_{x,y,z-\Delta z}(t - \Delta t) \quad (A.25)$$

Negative direction boundary condition:

$$u_{x,y,z+\frac{\Delta z}{2}}^z \left(t + \frac{\Delta t}{2} \right) = \frac{R_z - Z}{R_z + Z} u_{x,y,z+\frac{\Delta z}{2}}^z \left(t - \frac{\Delta t}{2} \right) - \frac{2}{R_z + Z} p_{x,y,z-\Delta z}(t - \Delta t) \quad (A.26)$$

A.3 Non-rectangular topology example mask set

This section gives the full set of masks for the non-rectangular topology example presented in Section 3.2.6. White boxes represent grid elements included in the simulated space while gray boxes are elements excluded from the simulation.

A.3.1 User-defined pressure element grid mask (pressure updates)

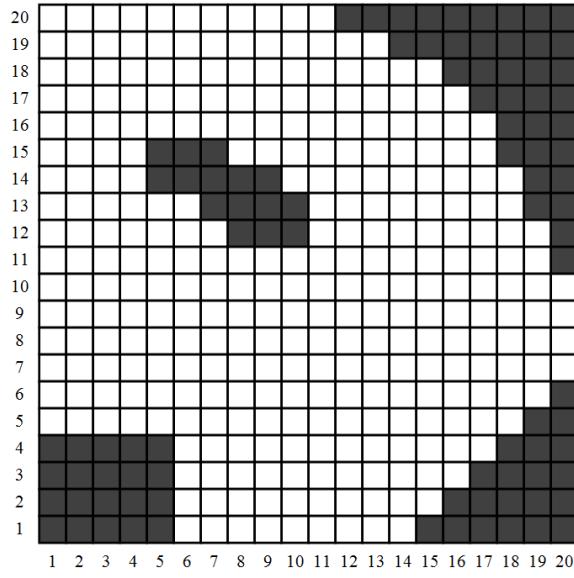


Fig. A.1 Pressure element grid mask, p_M

A.3.2 Non-boundary particle velocity element grid masks (velocity updates)

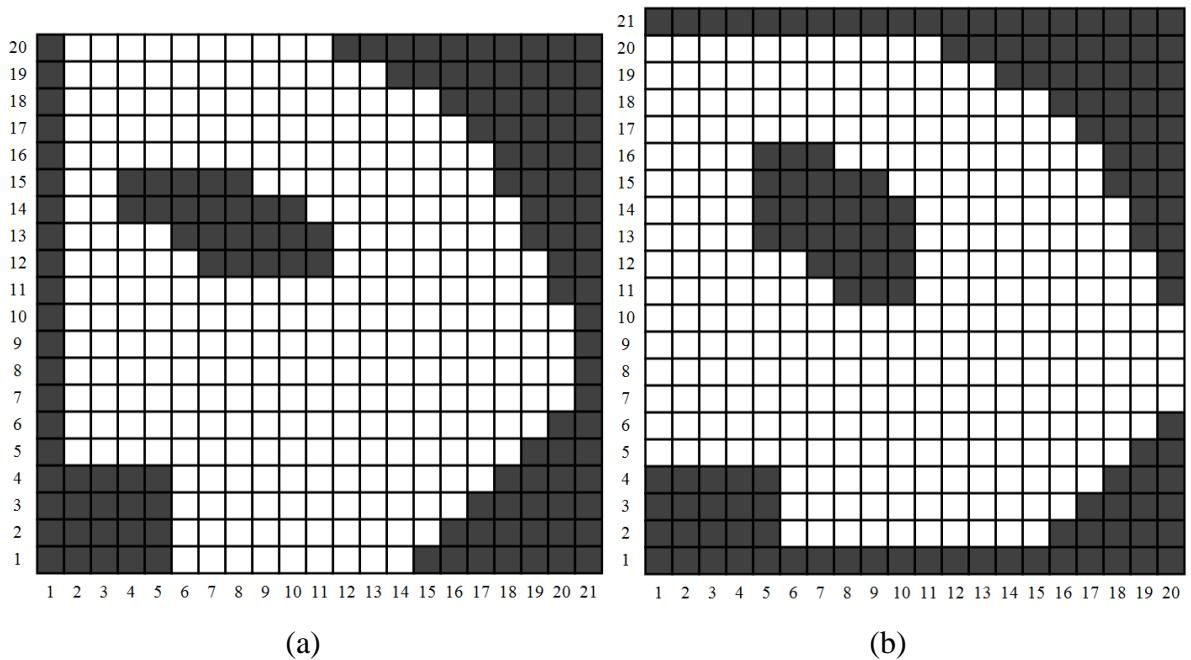


Fig. A.2 Non-boundary (a) x- and (b) y-dimension particle velocity grid masks, u_{xM} and u_{yM}

A.3.3 Boundary particle velocity element grid masks (velocity updates)

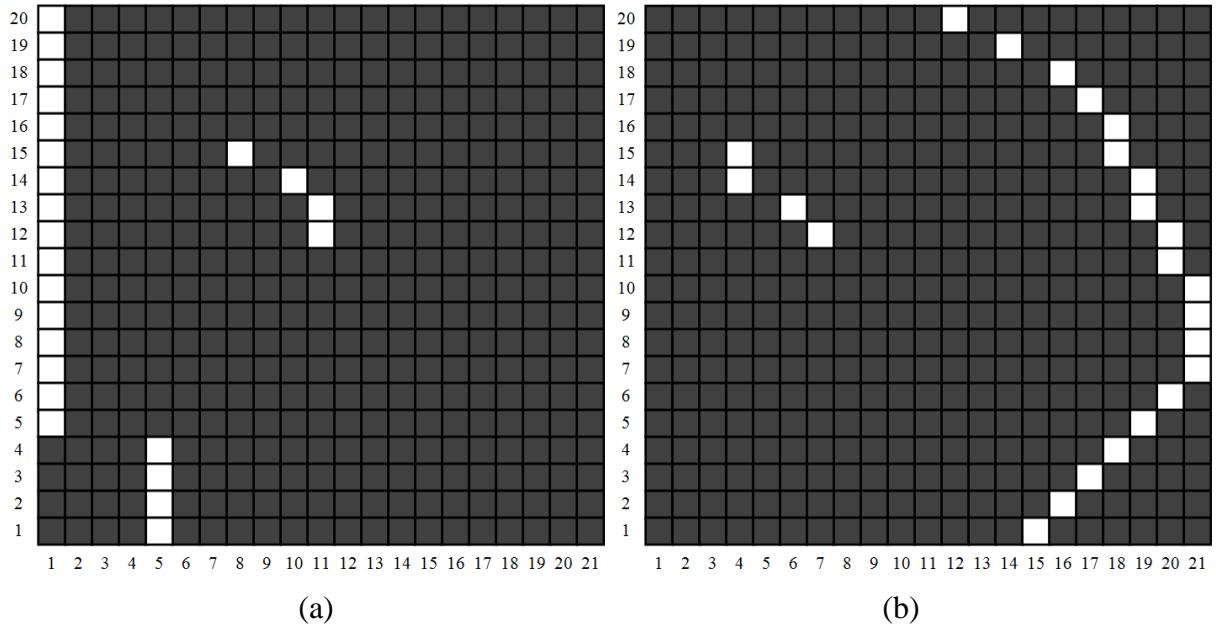


Fig. A.3 X-dimension boundary condition particle velocity masks for
 (a) negative x-direction direction, u_{ML} and u_{MR}

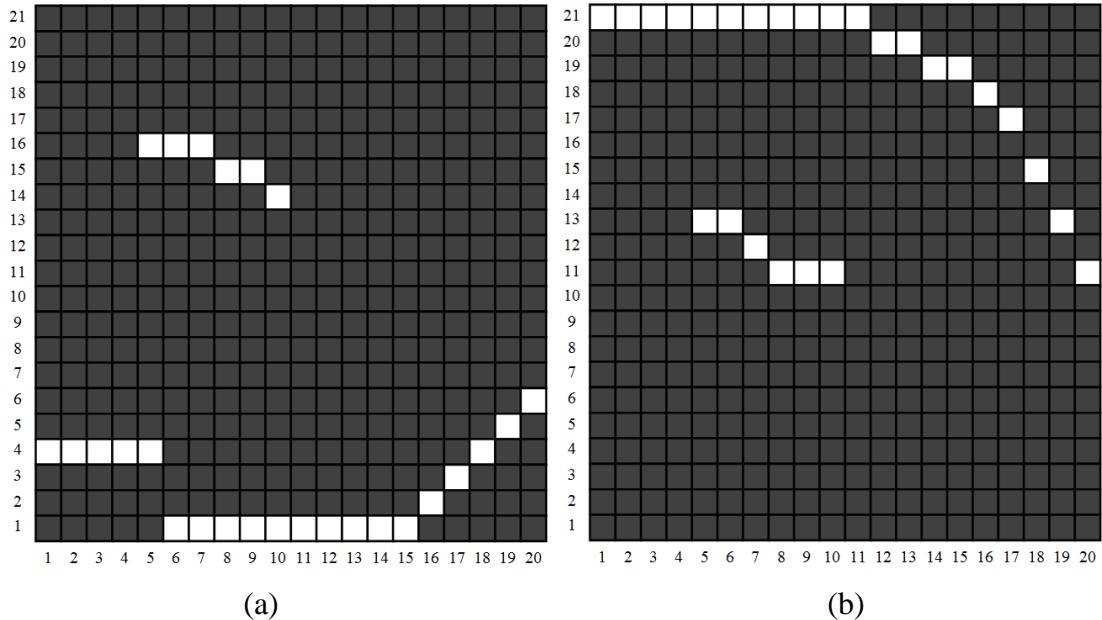


Fig. A.4 Y-dimension boundary condition particle velocity masks for
 (a) negative y-direction direction, u_{MB} and u_{MT}

A.3.4 Pressure element grid masks (non-boundary velocity updates)

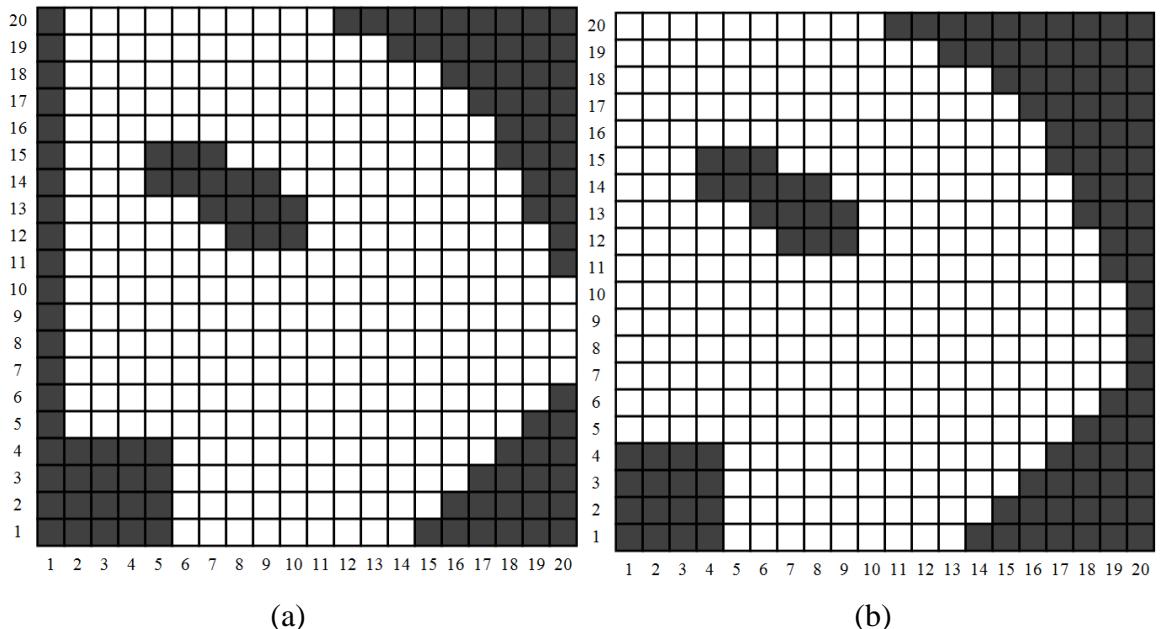


Fig. A.5 X-dimension non-boundary particle velocity update pressure grid masks for
(a) negative x-direction direction and (b) positive x-direction, p_{Mx1} and p_{Mx2}

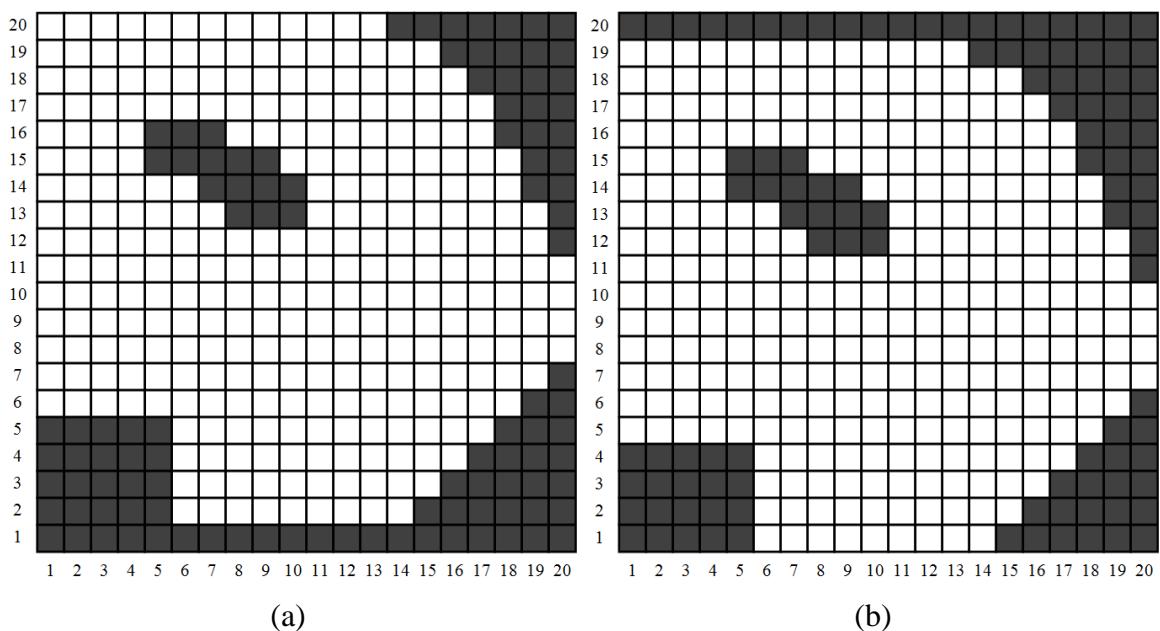


Fig. A.6 Y-dimension non-boundary particle velocity update pressure grid masks for
(a) negative y-direction direction and (b) positive y-direction, p_{My1} and p_{My2}

A.3.5 Pressure element grid masks (boundary velocity updates)

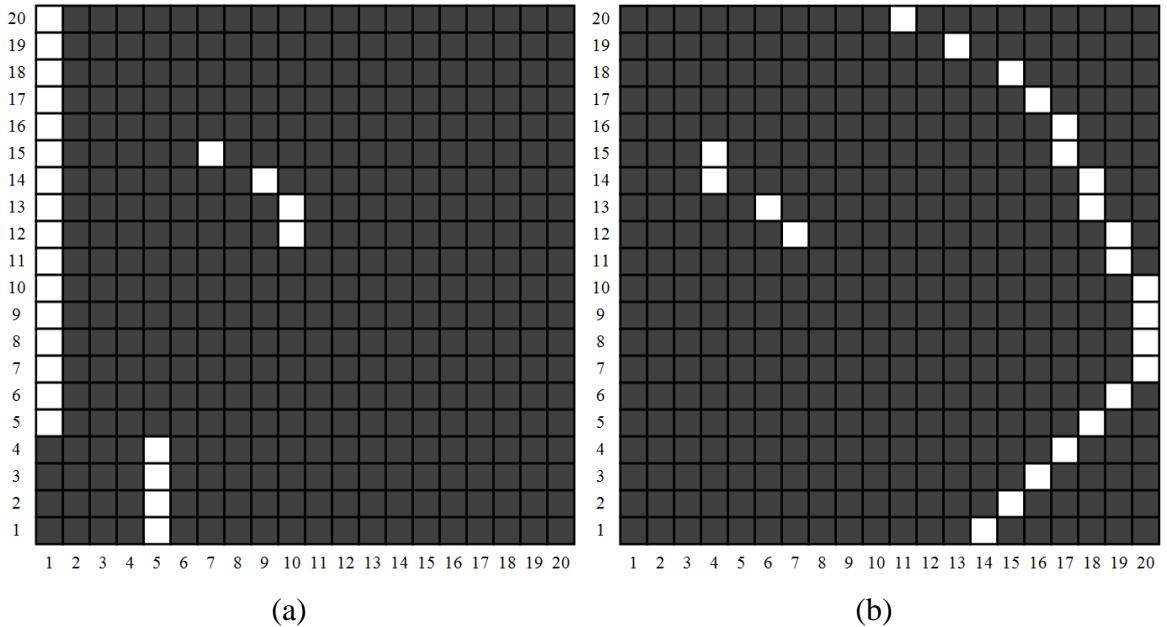


Fig. A.7 X-dimension boundary particle velocity update pressure grid masks for
 (a) negative x-direction direction and (b) positive x-direction, p_{ML} and p_{MR}

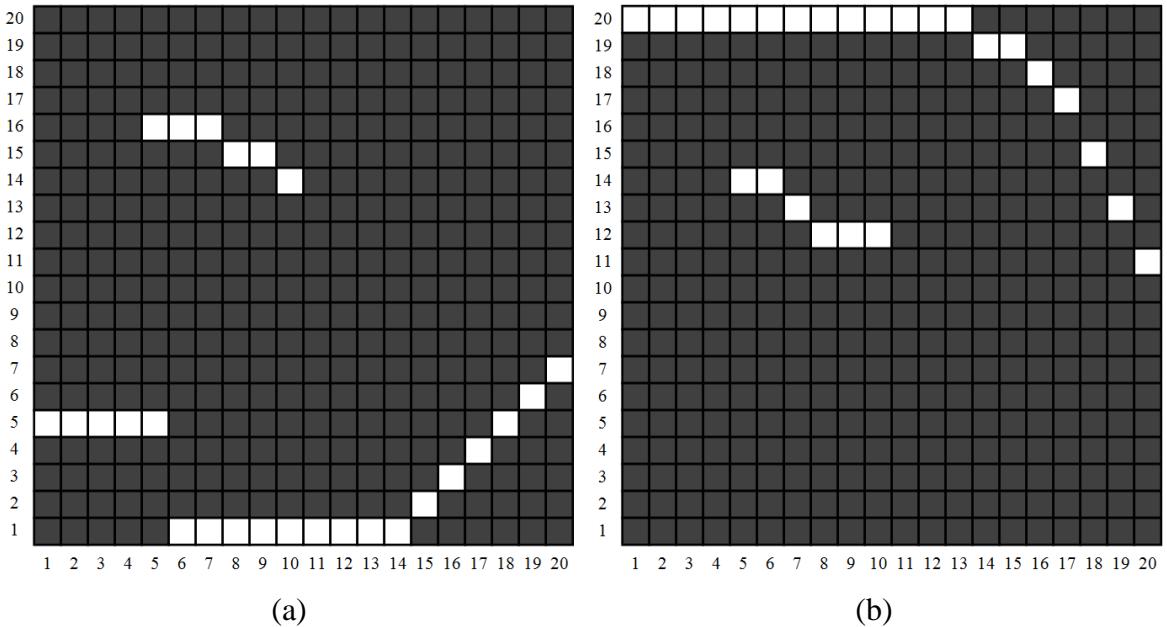


Fig. A.8 Y-dimension boundary particle velocity update pressure grid masks for
 (a) negative y-direction direction and (b) positive y-direction, p_{MB} and p_{MT}

A.3.6 Velocity element grid masks (pressure updates)

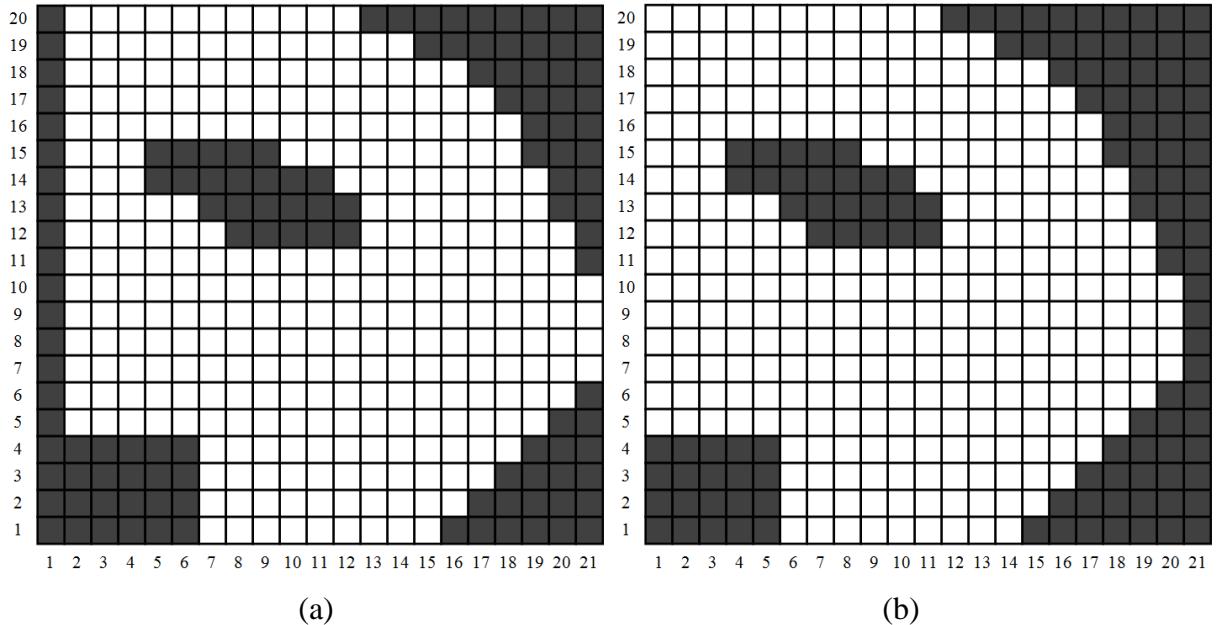


Fig. A.9 Pressure update x-dimension particle velocity grid masks for
(a) negative x-direction direction and (b) positive x-direction, u_{Mxp1} and u_{Mxp2}

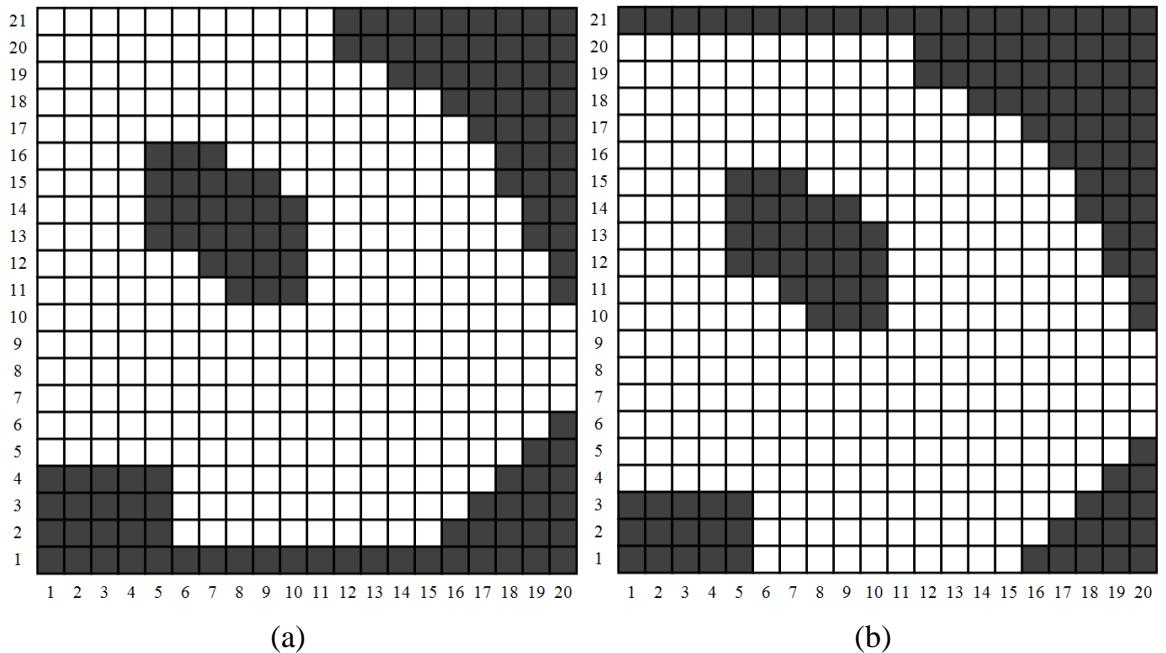


Fig. A.10 Pressure update y-dimension particle velocity grid masks for
(a) negative y-direction direction and (b) positive y-direction, u_{Myp1} and u_{Myp2}

Appendix B – FDTD simulation toolbox examples

This appendix contains a detailed discussion concerning the functionality of the FDTD toolbox, introduced in Chapter 3. The appendix first highlights the simulation configuration process and options as well as the data analysis possibilities. This discussion is followed by a series of worked examples which accentuate the powerful routines contained within the toolbox as they pertain to acoustics and audio engineering.

B.1 Toolbox features – configuration

Before any simulations can be carried out the virtual environment must be defined, both in the acoustical and electrical domain. All variables concerning the simulation configuration can be directly adjusted from within the GUI, allowing users to modify the model as they wish. This section gives an overview of the configuration sections, which allow the simulation variables to be properly adjusted while ensuring system stability is maintained.

B.1.1. Basic topology definition

The first task concerning the simulation configuration calls for a definition of the basic topographical properties. All other simulation parameters are disabled until these variables have been set. The acoustical space is addressed within the *Room Settings* box (Fig. B.1).

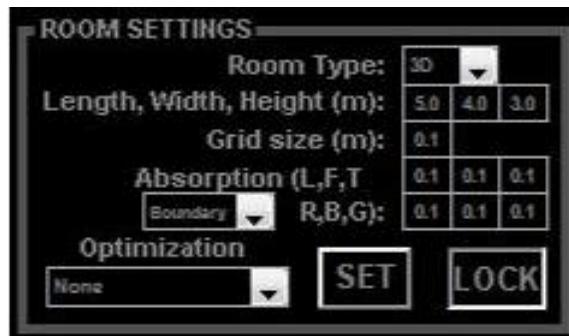


Fig. B.1 Layout of the *Room Settings* section

The room type dropdown menu permits toggling between two- and three-dimensional spaces, by switching between the relevant set of FDTD equations, as detailed in Chapter 3 and Appendix A, and enabling/disabling the third dimensional parameters concerning height and absorption. The length, width and height of the initial rectangular space are defined as any positive number greater than the grid size. The grid size, likewise, must be less than the smallest dimension of the core space. System stability is ensured by updating the sample rate of the simulation based on the grid size.

Two separate sets of frequency-independent absorption coefficients can be manipulated in this section; one for room boundaries and one for room obstacles. Absorption can range from 0 (no absorption, fully reflective surfaces) to 1 (full absorption, anechoic). Each wall and side of internal room obstacles has independently controllable absorption levels, allowing for sufficient control of the acoustical characteristics.

The *Room Settings* section also gives the choice of optimization routines aimed to assist in the design of an acoustical space and its corresponding sound reproduction system. The first optimization option concentrates on the rectangular room dimensions of the space, allowing a

sweep range to be defined for each primary rectangular dimension of the room. Once the remaining simulation parameters have been set, the simulation is run and repeated for each sweep point, while tracking the spatial variance over the defined listening locations.

Upon completion of the simulation sweep, spatial variance as a function of room dimensions is plotted (Fig. B.2) along with a text file which holds the data for each sweep point and identifies the best and worst dimensional configurations. This routine is particularly useful when designing a room with the goal of delivering the best possible low-frequency acoustical response across a wide listening area.

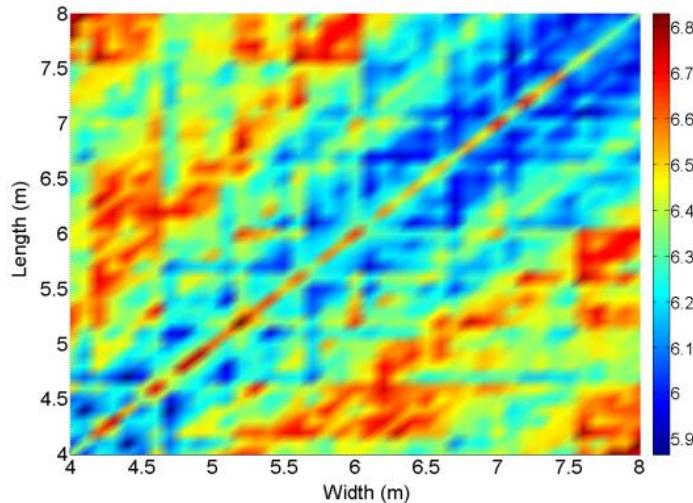


Fig. B.2 Example room dimension optimization results for a rectangular room of dimensions (3.5 – 10.5 m) x (3.5 – 10.5 m) x 3.5 m with a single omnidirectional subwoofer located at a room corner, plotting spatial variance (dB) as a function of room dimensions

In addition to room dimension optimization, there is an option for source position optimization. The routine operates in a similar manner to the dimension function, but instead of sweeping dimensions, one or more source position is swept across a defined area. Again, spatial variance is used as a metric for judging performance concerning the various configurations (Fig. B.3).

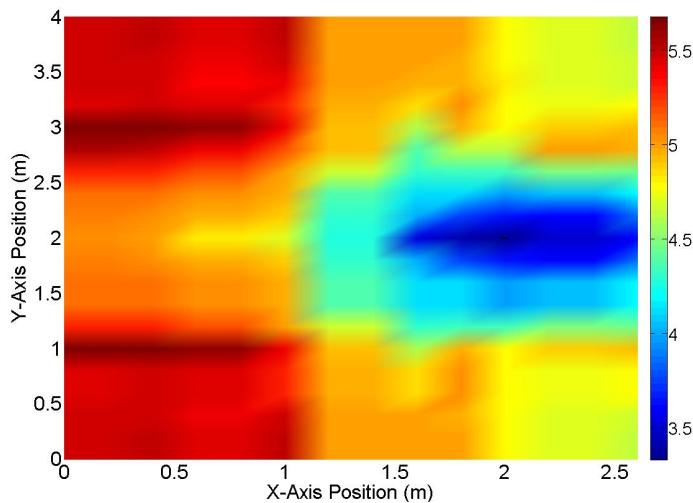


Fig. B.3 Example source position optimization results for a 5 m x 4 m x 3 m rectangular room with a single omnidirectional subwoofer (subwoofer height = 0.4 m), plotting spatial variance (dB) as a function of source position

Lastly, a live sound subwoofer optimization function is available which allows for a subwoofer system consisting of industry-standard cardioid subwoofers to be configured giving maximally even audience coverage while simultaneously minimizing the sound pressure level on stage. The sweep parameters include: subwoofer spacing, orientation, electronic time delay and polarity.

B.1.2. Spatial configuration

Once the basic acoustical space has been defined in the *Room Settings* section, the space can be modified to model non-rectangular topologies including obstacles within the room. The underlying methodology utilized for this process has been detailed in Chapter 3.2.6, and is controlled within the *Room Configuration* section (Fig. B.4).

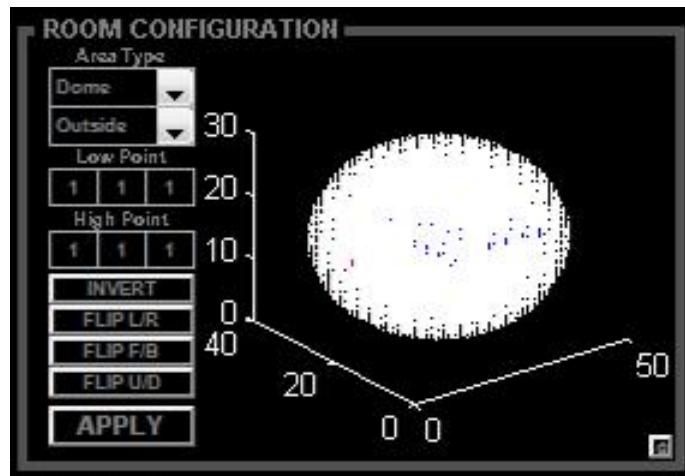


Fig. B.4 Layout of the *Room Configuration* section

The spatial configuration operates utilizing a “cookie-cutter” approach, whereby a user specifies a range of grid points to highlight using the *Low Point* and *High Point* index boxes. The selected grid points are shaped in a number of different ways, depending whether the topology is two- or three-dimensional. Two-dimensional topologies allow for three types of grid selection shapes: box, diagonal and curve. Three-dimensional topologies have similar shapes available, but with a third dimension, as well as additional shapes including a pyramid and a dome. These shapes can be rotated in any direction as well as inverted. Once the grid points have been selected, the user can apply them to the pressure point grid which updates the remaining grid set (Fig. B.5).

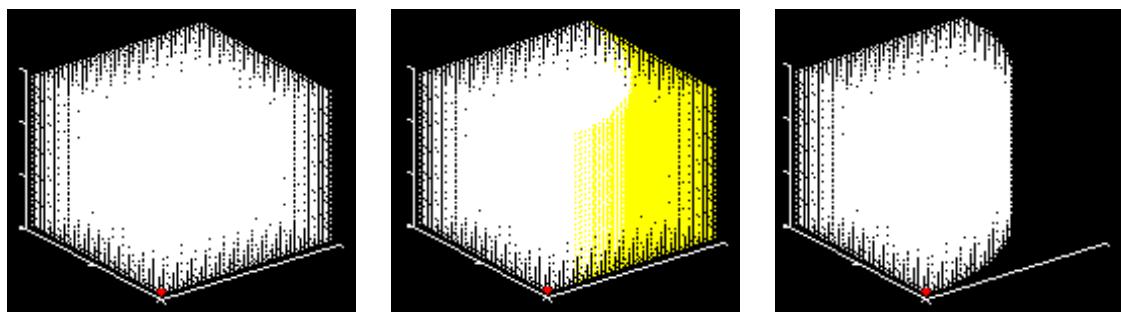


Fig. B.5 Non-rectangular topology adjustments in the *Room Configuration* section

B.1.3. Source setup

The sources are placed within the model using the *Source Settings* section (Fig. B.6).

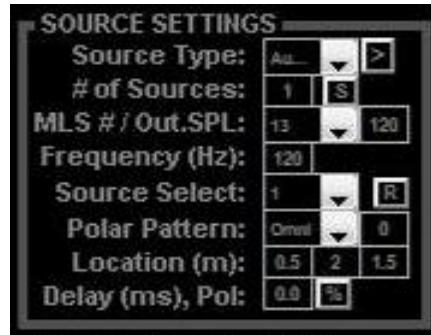


Fig. B.6 Layout of the *Source Settings* section

There can be nearly any number of sources in the simulation model, only limited by the amount of computer memory, as each source has a distinct input signal, which must be held in the computer's RAM during simulation. The practical upper limit for number of sources is therefore between one and two thousand. Most realistic applications, of course, only require one to eight sources.

Each source has a number of individually controllable parameters including: position, polar pattern, electronic delay and polarity. Source positions are tracked within the *sgrid* variable in the code (see Fig. 3.9, line 33) allowing for the sources to be specifically addressed during the simulation. The available polar patterns are omnidirectional, dipolar and cardioid. These patterns can be rotated over a full 360° , with 30° resolution. The electronic delay and polarity specifications for each source are applied by the source signal generation function, which is discussed in the next section of this chapter.

Finally, the signal type for all sources can be defined within this section. Choices include: Gaussian pulse, sinusoid, maximum length sequence (MLS), swept sine, tone bursts, tone burst sweep and any .wav audio file. Depending on the source signal choice, there will be various adjustable signal parameters to give maximal flexibility concerning the source signal. The output amplitude can be set (in dB) which is the same for all sources. This does not affect the output results of the simulation, as linearity is assumed, but can assist to scale the simulation amplitude to make the results relevant to reality.

B.1.4. Signal generation

Care must be taken when generating source input signals. It is important that the user-defined delay, polarity and amplitude are all properly handled within the signal generation routine. All signal generation is handled within *sourceGen.m*, with the specified simulation details passed from the main function *simGUI.m*. The signals are generated only after the *Run Simulation* button has been pressed in the GUI.

First, the required electronic delay values are examined to determine the maximum delay value, thus dictating the required signal length. The signals are then generated for each source, one at a time, in an appropriate section depending on the source type.

Gaussian pulses are generated with Eq. B.1, where t_0 is the specified electronic delay and w is the width of the pulse. The signal is scaled to match the required amplitude, A .

$$y = A e^{-0.5 \left(\frac{t-t_0}{w} \right)^2} \quad (B.1)$$

Sinusoidal signals are generated in a similar manner, but with the signal shifted in time according to the required delay. This is performed after the raw sinusoid is generated (Eq. B.2) where f is the sinusoidal frequency.

$$y = A \sin(2\pi f t) \quad (B.2)$$

MLS signals are generated by calling *mlsGen.m*. This function generates the MLS signal based on the recursion relationship for the specified MLS order, m (Fig. B.7). The MLS signal is repeated once, in accordance with MLS measurement procedures detailed in the Chapter 3, and then delayed, as appropriate.

```

79      % Determine length of MLS
80 -    N = (2^m) - 1;
81
82      % Initialize MLS output vector
83 -    mls = zeros(N, 1);
84
85      % Create and initialize the shift register (initial condition all ones)
86 -    reg = ones(m, 1);
87
88      % Loop to create the MLS sequence and update the shift register
89 -    for n = 1 : N
90        % Set the proper point in the output sequence
91 -        mls(n) = reg(1);
92        % Determine the new value for the mth element in the shift register
93 -        temp = mod(sum(reg(rel)), 2);
94        % shift the values in the register
95 -        reg(1 : m - 1) = reg(2 : m);
96 -        reg(m) = temp;
97 -    end
98
99      % shift the scale from 0 to 1 to -1 to 1
100 -   mls = (-1).^mls;
```

Fig. B.7 Code for MLS signal generation (rel = recursion relationship)

Swept sinusoids are generated based on the user-defined low and high frequency values. The generated signal sweeps the frequency range logarithmically, as specified in Eq. B.3, where f_{min} and f_{max} are the start and stop frequencies, respectively, and T is the signal duration (s).

$$y(n) = A \sin \left(\frac{2\pi f_{min} T}{\log \left(\frac{f_{max}}{f_{min}} \right)} \left(e^{\frac{t(n)}{T} \log \left(\frac{f_{max}}{f_{min}} \right)} - 1 \right) \right) \quad (B.3)$$

Tone bursts are generated based on three user-defined variables: burst frequency, f , cycles per burst, *cycles*, and burst repetitions, *reps*. First, a pure sinusoidal signal is generated of length corresponding to the total required burst time (Eq. B.4). Next a window must be generated which forms the shape of the tone bursts. In this application, a raised cosine is implemented (Eq. B.5).

$$y_{single} = A \sin(2\pi ft) \quad (B.4)$$

$$w = 0.5 - 0.5 \cos\left(\frac{2\pi ft}{cycles}\right) \quad (B.5)$$

The window is applied to the pure sinusoid to give a single repetition of the burst. If more than one repetition is required, the signal is repeated, as shown in Fig. B.8.

```

102 % calculate length of single tone burst
103 - yLen = length(y2);
104
105 % add additional repetitions to the final signal
106 - for r = 1 : reps
107     % determine index of next repetition
108     i1 = yLen*(r-1) + 1;
109     % repeat single tone burst at the end of the signal
110     y(i1 : i1 + yLen - 1) = y2;
111 - end

```

Fig. B.8 Code for adding additional tone burst repetitions ($y_2 = \text{single tone burst}$)

Lastly, audio signals from .wav files must be handled with care to avoid any spectral aliasing due to sample rate mismatch between the source signal and simulation. Since the FDTD simulation is only accurate up to a frequency defined by the grid spacing, the audio file must pass through a complementary crossover network, where the low-band output is fed into the simulation and the high-band output is used for auralization purposes (detailed later in this appendix).

The complementary crossover network is generated using a set of finite impulse response (FIR) filters. The low-pass filter frequency response is initialized to one, and then all frequency bins above the crossover frequency are set to zero. Next, an all-pass FIR filter is generated and the low-pass filter coefficients are subtracted from the all-pass filter to give the complementary high-pass filter (Fig. B.9).

```

19 % initialize the frequency magnitudes and frequency vector
20 - f = linspace(0, 1, N-1);
21 - m1 = ones(1, N-1);
22
23 % determine the normalized crossover frequency
24 - nfc = fc/(Fs/2);
25
26 % locate the crossover point in the frequency & magnitude vectors
27 - ind = find(f > nfc, 1, 'first');
28
29 % set the frequencies above the crossover point to zero
30 - m1(ind : end) = 0;
31
32 % create the low pass filter
33 - bL = fir2(N, f, m1, chebwin(N+1, 120));
34
35 % create the allpass filter
36 - m2 = ones(1, N-1);
37 - bA = fir2(N, f, m2, chebwin(N+1, 120));
38
39 % create the high pass filter
40 - bH = bA - bL;

```

Fig. B.9 Code for crossover network complementary FIR filter generation

B.1.5. Virtual listener setup

The final configuration procedure is to label certain pressure grid points as listening locations. These locations record the received pressure levels over the duration of the simulation for post-simulation data analysis. Listening location positioning is adjusted in the *Measurement Settings* section (Fig. B.10).



Fig. B.10 Layout of the *Measurement Settings* section

By default, listening locations are organized in a grid-structure, allowing the number of locations to be the square of one to ten (i.e. 1, 4, 9, 16, etc.). Once the required number of listening location has been chosen, the location of the grid center point is defined along with the point spacing (tight, normal or loose). Alternatively, the *M* button can be pressed, allowing for manual positioning of any listening point.

All listening locations are assigned a unique color which corresponds to the plots generated in the post-simulation data analysis and also in the *Auralization* section.

B.2 Toolbox features - simulation

A number of options are built into to FDTD simulation toolbox providing a number of useful data animation/visualization functions as the simulation is running and immediately after it has completed. The animation and visualization options are specified in a dropdown menu in the *Simulation Settings* section (Fig. B.11).



Fig. B.11 Layout of the *Simulation Settings* section

B.2.1. Animations

A useful tool when researching room acoustics is a method to animate sound wave propagation across a virtual space. This is made available in the toolbox and functions simultaneously with the FDTD simulation procedure. The pressure grid data can be animated either in two or three dimensions (Fig. B.12).

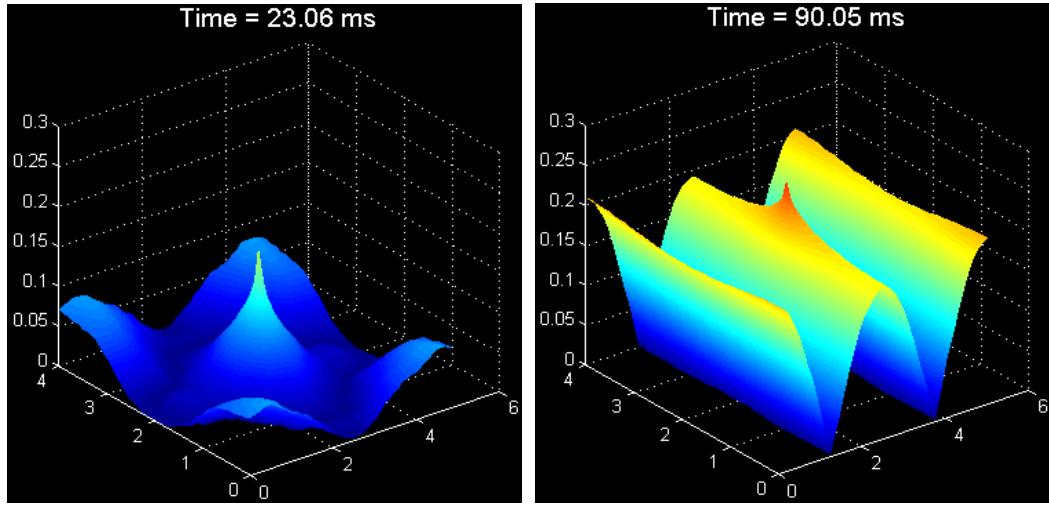


Fig. B.12 Example of the animation function for a two-dimensional topology illustrating the buildup of a standing wave pattern due to a single source at the room center

If the simulated space is two-dimensional, the *2D Animation* menu option plots the data in a two-dimensional contour plot. If the *3D Animation* option if chosen, the data is plotted in a 3D contour plot where the vertical axis is the absolute pressure amplitude (Figs. B.13 & B.14).

```

380 % plot pressure grid point data for animation
381 surf(linex, liney, abs(p));
382 shading interp;
383 % determine to display time or frequency values
384 if sType == 4 % frequency sweep
385 title(sprintf('Frequency = %.2f Hz',fVals(n)),...
386 'Color',[1 1 1],'FontSize', 14);
387 else % standard time domain
388 title(sprintf('Time = %.2f ms',n*dt*1e3),...
389 'Color',[1 1 1],'FontSize', 14);
390 end
391 xlabel('Width (meters)', 'Color', [1 1 1]);
392 ylabel('Length (meters)', 'Color', [1 1 1]);
393 view(2); % switch to 2D view
394 axis tight;
395 caxis([0 pS]); % adjust color scale
396 drawnow;

```

Fig. B.13 Code for *2D Animation* option for two-dimensional topologies

```

447 surf(linex, liney ,abs(p));
448 shading interp;
449 lighting phong;
450 % determine to display time or frequency values
451 if sType == 4 % frequency sweep
452 title(sprintf('Frequency = %.2f Hz',fVals(n)),...
453 'Color',[1 1 1],'FontSize', 14);
454 else % standard time domain
455 title(sprintf('Time = %.2f ms',n*dt*1e3),...
456 'Color',[1 1 1],'FontSize', 14);
457 end
458 xlabel('Width (meters)', 'Color', [1 1 1]);
459 ylabel('Length (meters)', 'Color', [1 1 1]);
460 zlabel('Pressure (linear)', 'Color', [1 1 1]);
461 axis tight;
462 caxis([0 pS]); % adjust color scale
463 zlim([0 pS]); % adjust vertical axis limits
464 drawnow;

```

Fig. B.14 Code for *3D Animation* option for two-dimensional topologies

Alternatively, if the simulated space is three-dimensional, the *2D Animation* option plots the data only in the horizontal plane at a height specified by the *z-Slice* text box in the GUI. The *3D Animation* option plots a three-dimensional grid structure representing all pressure points (Figs. B.15 & B.16).

```

526 % plot pressure grid point data for animation
527 - surf(linex, liney, abs(p(:,:,zG)));
528 - shading interp;
529 % determine to display time or frequency values
530 - if sType == 4 % frequency sweep
531 -     title(sprintf('Frequency = %.2f Hz',fVals(n)),...
532 -         'Color',[1 1 1],'FontSize', 14);
533 - else % standard time domain
534 -     title(sprintf('Time = %.2f ms',n*dt*1e3),...
535 -         'Color',[1 1 1],'FontSize', 14);
536 - end
537 - xlabel('Width (meters)', 'Color', [1 1 1]);
538 - ylabel('Length (meters)', 'Color', [1 1 1]);
539 - view(2); % switch to 2D view
540 - axis tight;
541 - caxis([0 pS/5]);% adjust color scale
542 - drawnow;
```

Fig. B.15 Code for *2D Animation* option for three-dimensional topologies

```

587 % loop to convert the pressure grid matrix to a vector
588 - pVec = reshape(p, 1, size(p, 1)*size(p, 2)*size(p, 3));
589 - [row, col] = find(~isnan(pVec));
590 % plot the scatter plot of the pressure data
591 - scatter3(col1,row1,hgt1,8,pVec(col), 'Marker', 's');
592 - axis tight;
593 - caxis([0 pS/5]);% adjust color scale
594 - alpha(0.4); % adjust grid point transparency
595 - if sType == 4 % frequency sweep
596 -     title(sprintf('Frequency = %.2f Hz',fVals(n)),...
597 -         'Color',[1 1 1],'FontSize', 14);
598 - else % standard time domain
599 -     title(sprintf('Time = %.2f ms',n*dt*1e3),...
600 -         'Color',[1 1 1],'FontSize', 14);
601 - end
602 - xlabel('Width (meters)', 'Color', [1 1 1]);
603 - ylabel('Length (meters)', 'Color', [1 1 1]);
604 - zlabel('Height (meters)', 'Color', [1 1 1]);
605 - drawnow;
```

Fig. B.16 Code for *3D Animation* option for three-dimensional topologies

B.2.2. Final spatio-pressure distribution plots

Since the animation functions can cause significant increases to simulation runtime, an alternative visualization option is available which plots the final spatio-pressure distribution. This allows the simulation to run without the added processing requirements of continuously updating the animation plot. The final values in the pressure grid data matrix are plotted immediately following the conclusion of the simulation. This gives users the ability to immediately examine any modal distribution problems in the virtual acoustical space.

As with the animations, there exist the same two- and three-dimensional plotting options, with the pressure values are plotted on a logarithmic scale (Fig. B.17).

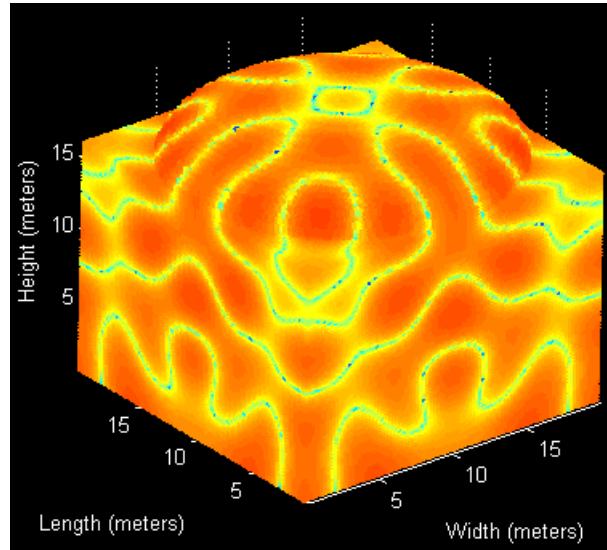


Fig. B.17 Example of the final SPL plotting function for a three-dimensional space

B.2.3. Auralization

An auralization function is built into the toolbox available immediately after the simulation is finished. This allows users to listen to sounds received at the listening locations and compare how they differ subjectively from point to point. A set of colored boxes appears in the *Auralization* section upon completion of the simulation, related to the location colors shown in the room layout plots (which is discussed in the following section) (Fig. B.18).

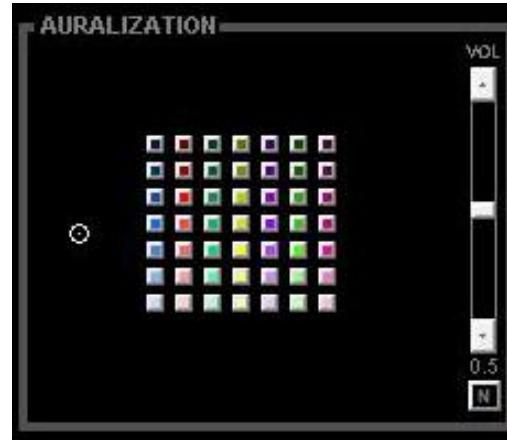


Fig. B.18 Layout of the *Auralization* section

The simulation results are band-limited due to the limited sample rate (based on the grid spacing) to ensure system stability. As noted with audio files, the input signals must pass through a complementary crossover network to avoid spectral aliasing. To provide a full-range auralization, the two output bands of the crossover network must be recombined.

The first step in this process involves converting the simulation output to the original file sample rate. This is accomplished with the *resample* function in MATLAB, which has anti-aliasing methods built into its routine. Once the sample rate has been restored to its original value, the high frequency band, which was not passed through the simulation, must be delayed. The delay is calculated for each listening location and added to the beginning of the high frequency signal. The two signals are finally combined for the auralization.

B.3 Toolbox features – analysis

One valuable aspect of the toolbox is the direct data analysis capabilities. These functions were central in the progression of the work presented in this thesis and expedited research in a number of key topic areas. All analysis options are accessed within the *Analysis* and *Analysis Results* sections (Fig. B.19). The following subsections highlight the key analysis options built into the toolbox.

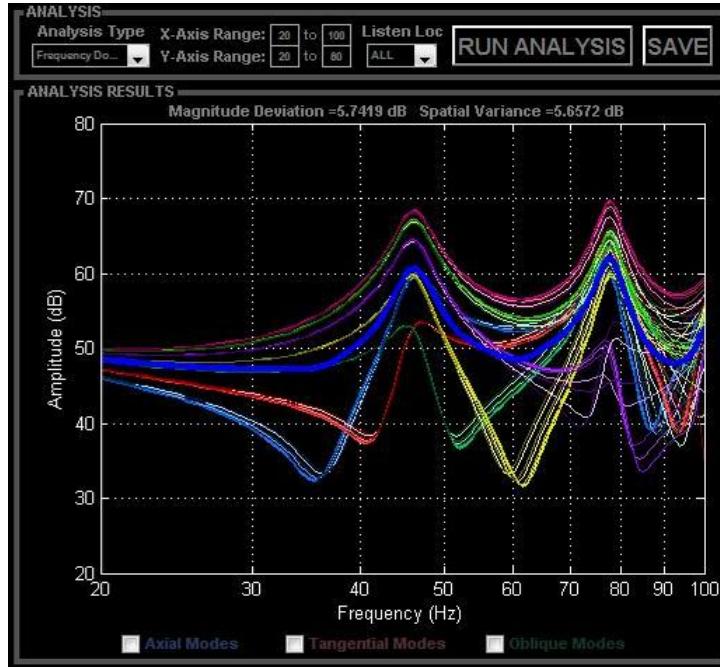


Fig. B.19 Layout of the *Analysis* and *Analysis Results* sections

B.3.1. Room layout

The first analysis option, available both before and after running the simulation, is the room layout function. This provides a floor plan of the current simulation configuration, including the sources and listeners (Fig. B.20). This is especially useful when constructing a complicated virtual space, where verifying a correct layout before running the simulation can save a considerable amount of time.

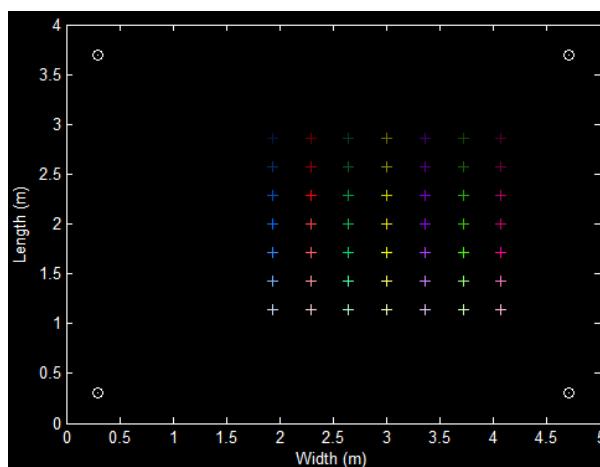


Fig. B.20 Example of a 5 m x 4 m room layout plot, including four subwoofers in the room corners and a 49-point rectangular virtual listener grid

B.3.2. Theoretical mode calculator

Another analysis option that is available both before and after running the simulation is a theoretical room-mode calculator. The first routine presents a plot of the modal frequencies and their number of occurrences up to a specified frequency in the *Analysis* section. This option is useful when choosing room dimensions as it shows spectral spacing of modal frequencies. As example, the modal spectral distribution for a cubic room of dimensions 5 m x 5 m x 5 m (Fig. B.21a) indicates multiple modes sharing the same frequency, which is expected to cause a high spatial variance at these frequencies. Alternatively, if the room dimensions are adjusted to 5 m x 4 m x 3 m, the spectral distribution shows less modal overlap, which likely result in less spatial variance in the listening area (Fig. B.21b).

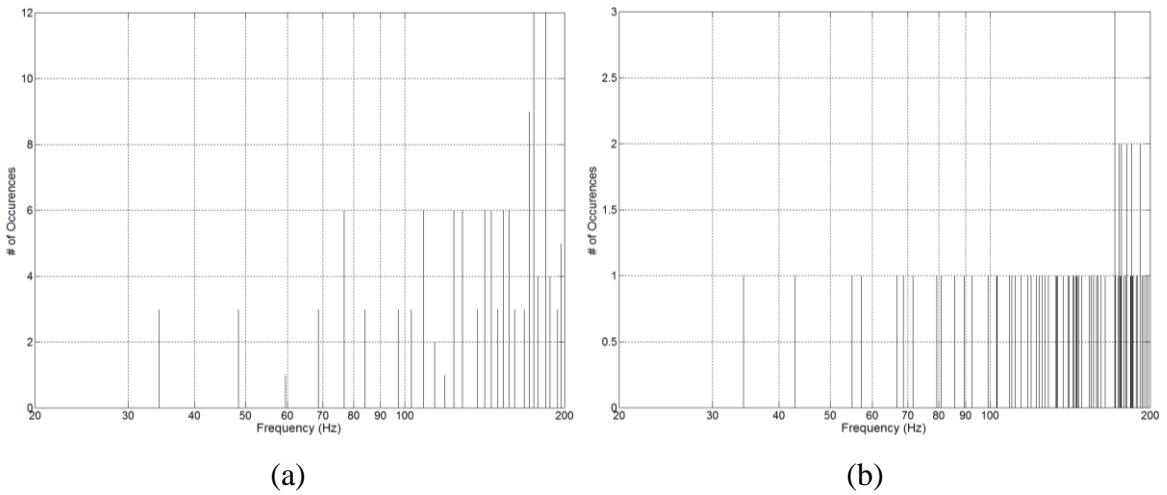


Fig. B.21 Example theoretical modal spectral distribution plots for (a) 5 m cubic room and (b) 5 m x 4 m x 3 m room

In addition to examining the modal spectral distribution, the toolbox provides an option for plotting modal spatial distribution. This operates by calculating theoretical nodal and antinodal planes/lines/points in the virtual space. Axial modes consist of nodal/antinodal planes, spaced along the dimension of a given mode. The planes are separated by the quarter-wavelength of the modal frequency (Fig. B.22a), whereby the first and last antinodal plane is located at room boundaries (Fig. B.22b).

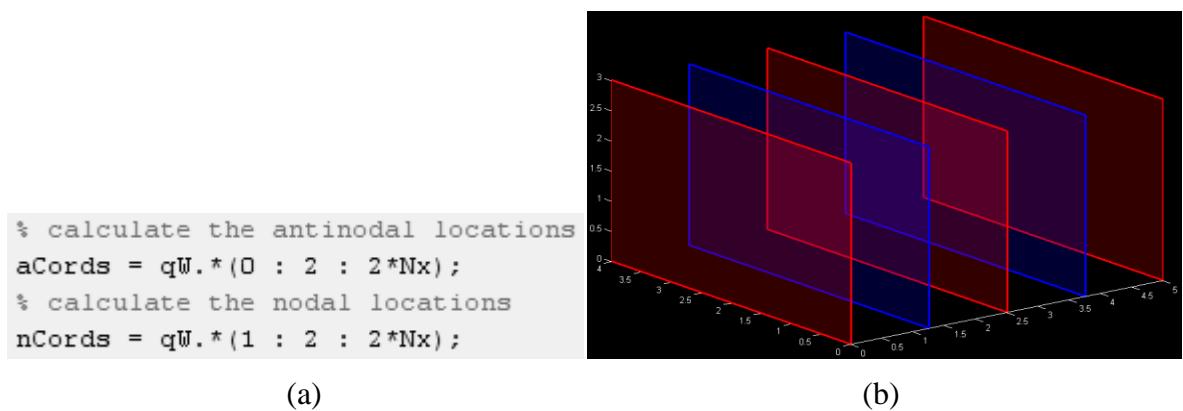


Fig. B.22 Example axial mode spatial distribution calculation (a) and distribution plot (b) (red = antinode, blue = node, qW = quarter wavelength (m), Nx = non-zero mode number)

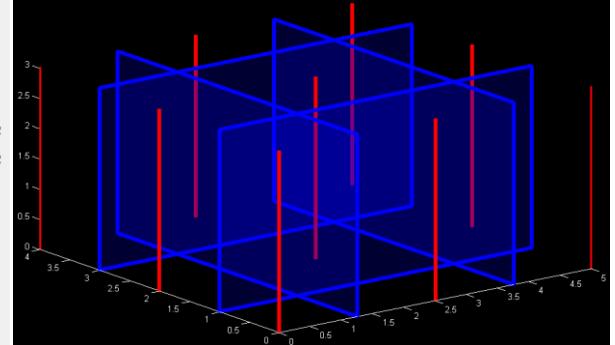
Similarly, spatial distribution plots can be generated for tangential and oblique modes. Tangential modes have nodal planes and antinodal lines, as calculated in Fig. B.23a and plotted

in Fig. B.23b. Instead of simple quarter-wavelength spacing, as with axial modes, tangential modal spacing is calculated based on the room dimensions and mode numbers to give even nodal/antinodal spacing across each dimension.

```
% determine the proper step sizes
qWx = lx/(2*Nx);
qWy = ly/(2*Ny);

% calculate the locations of the antinodes
aCords = zeros(Nx + 1, Ny + 1, 2);
aCords(:, 1, 1) = qWx.*([0 : 2 : 2*Nx];
aCords(1, :, 2) = qWy.*([0 : 2 : 2*Ny];
for x = 2 : Nx + 1
    for y = 2 : Ny + 1
        aCords(x, y, 1) = aCords(x, 1, 1);
        aCords(x, y, 2) = aCords(1, y, 2);
    end
end

% calculate the locations of the nodes
nCordsX = qWx.*([1 : 2 : 2*Nx];
nCordsY = qWy.*([1 : 2 : 2*Ny];
```



(a)

(b)

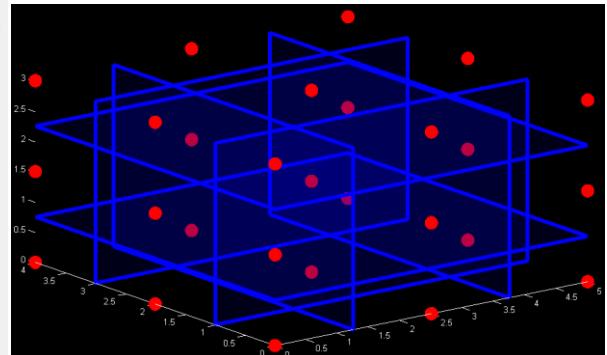
Fig. B.23 Example tangential spatial distribution calculation (a) and distribution plot (b) (red = antinode, blue = node, qW = quarter wavelength (m), Nx,y = non-zero mode numbers)

Oblique mode spatial distributions are calculated and plotted in a similar manner to tangential modes (Fig. B.24a), except with antinodal points and nodal planes (Fig. B.24b).

```
% determine the proper step sizes
qWx = lx/(2*Nx);
qWy = ly/(2*Ny);
qWz = lz/(2*Nz);

% calculate the locations of the antinodes
aCordsX = zeros(Nx + 1, Ny + 1, Nz + 1);
aCordsY = zeros(Nx + 1, Ny + 1, Nz + 1);
aCordsZ = zeros(Nx + 1, Ny + 1, Nz + 1);
for z = 1 : Nz + 1
    for y = 1 : Ny + 1
        aCordsX(:, y, z) = qWx.*([0 : 2 : 2*Nx];
    end
end
for z = 1 : Nz + 1
    for x = 1 : Nx + 1
        aCordsY(x, :, z) = qWy.*([0 : 2 : 2*Ny];
    end
end
for y = 1 : Ny + 1
    for x = 1 : Nx + 1
        aCordsZ(x, y, :) = qWz.*([0 : 2 : 2*Nz];
    end
end

% calculate the locations of the nodal planes
nCordsX = qWx.*([1 : 2 : 2*Nx];
nCordsY = qWy.*([1 : 2 : 2*Ny];
nCordsZ = qWz.*([1 : 2 : 2*Nz];
```



(a)

(b)

Fig. B.24 Example oblique spatial distribution calculation (a) and distribution plot (b) (red = antinode, blue = node, qW = quarter wavelength (m), Nx,y,z = mode numbers)

B.3.3. Time domain & frequency domain analysis

Once the simulation has completed, there are a number of data analysis options. The first two include time and frequency domain analysis. The time domain option plots the recorded data from each listening location directly, either on a linear or logarithmic scale. The frequency domain option takes the time domain data and applies a fast Fourier transform (FFT) to the data and plots the resulting frequency response on a logarithmic scale. An example of these options is shown in Fig. B.25, where a single subwoofer was measured at one listening location using a 120 Hz sinusoid source signal.

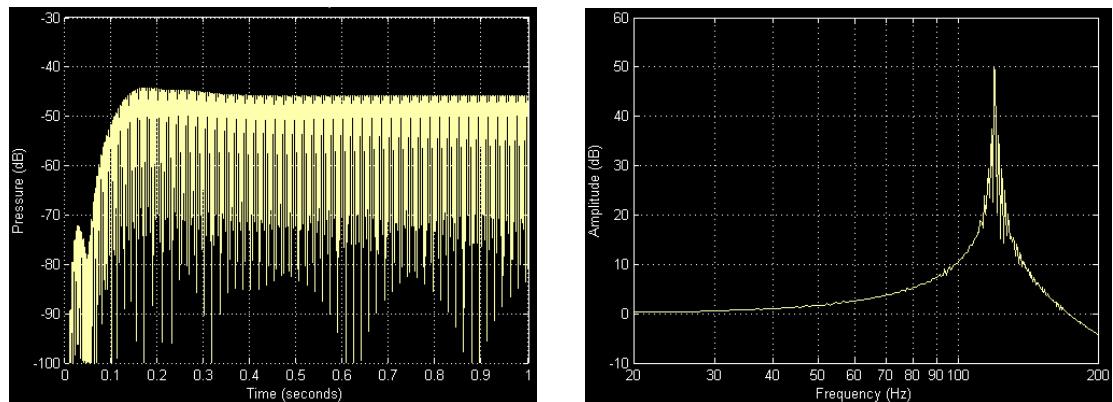


Fig. B.25 Examples of time (left) and frequency (right) domain analysis with an example simulated measurement of a 120 Hz sinusoid at a single listening location

B.3.4. Spectrogram

A useful alternative to the time/frequency domain methods is a spectrogram. This is useful when tracking the frequency response over a time period. The spectrogram is generated by taking FFTs at overlapping time windows in *spect.m*. The FFT data for each window is plotted as a contour plot using MATLAB's *surf* function, producing the spectrogram. As an example, a room's impulse response is calculated using an MLS measurement and plotted as a spectrogram at a single listening location (Fig. B.26).

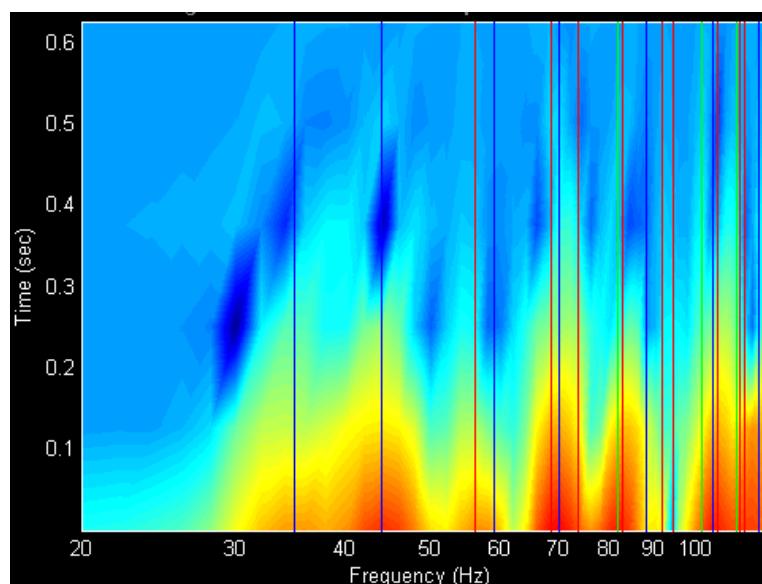


Fig. B.26 Example of a spectrogram for an MLS measurement in a 5 m x 4 m x 3 m room with theoretical axial (blue), tangential (red) and oblique (green) modes overlaid

B.3.5. Tone bursts

The first option for tone burst analysis is direct plotting from the measurements. The tone burst comparison option allows for three burst measurements to be compared side by side, along with the source signal, and also played back for subjective comparisons (Fig. B.27).

While single-frequency tone burst testing provides insight into transient behavior in a room, it is beneficial to have a method to combine data from numerous tone burst tests sweeping a given frequency range. This requires the reduction of data dimensions. The raw tone burst data at a single point and frequency contains two dimensions: time and amplitude. Expanding this to include all listening points adds two more dimensions corresponding to room location (three if in three dimensions). Performing multiple tone burst tests over a frequency range adds the final dimension, frequency. For a two-dimensional space this amounts to five dimensions. Obviously, this cannot be clearly represented in a single plot.

The simplest dimension compression is performed on the listening location dimensions. Each location is given a unique index number. These indices are arranged along one dimension with listening points in the same row in groups together.

The time dimension is compressed into a single point by comparing the measured signal to the source signal. This is accomplished by extracting the envelopes of both signals using a Hilbert transform. The absolute values of these envelopes are compared using a statistical analysis called mean absolute error (MAE) (Eq. B.6). MAE is the average deviation the measured waveform envelope exhibits from the source waveform envelope. When normalized signals are compared, the MAE will be between zero and one, with an MAE of zero indicating perfect reproduction of the source signal and an MAE of one indicating the measured signal is virtually unrecognizable to the source signal.

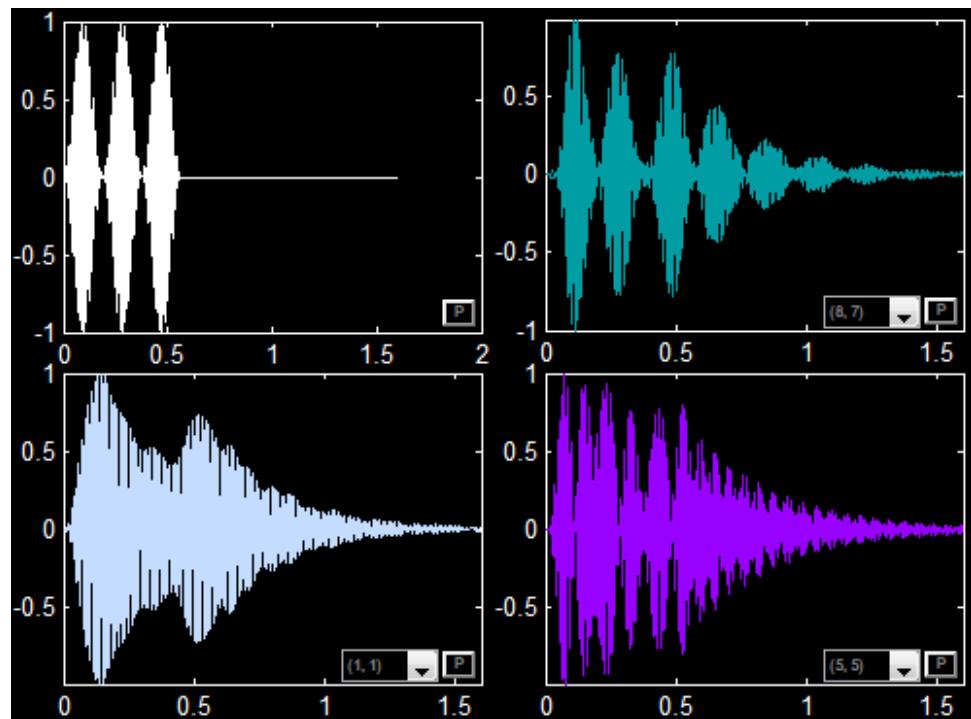


Fig. B.27 Example of the tone burst comparison option where the upper left hand plot is the source signal and the remaining three plots are distinct listening locations within the room

$$MAE = \frac{1}{N} \sum_{i=1}^N |s_i - y_i| \quad (B.6)$$

where: s_i = source signal at sample i
 y_i = measured signal at sample i
 N = total number of samples

This final reduction leaves three dimensions: listening location, frequency and MAE. These are easily represented where the x-axis is the listening location index, the y-axis is the tone burst frequency and the color (or z-axis) is the MAE. An example is given in Fig. B.28, where a 15-cycle, 5 repetition tone burst is measured for frequencies between 20 and 120 Hz with 2 Hz resolution.

It can be beneficial to rework the dimensional compression to give a plot that shows individual location response over all frequencies in a spatial layout. This is accomplished by taking the average MAE of all frequencies at a single listening location and adding the standard deviation of the MAE at that location. This gives an overall quality rating for each location over the tested frequency band. In essence, this indicates the prime listening locations as far as transient response is concerned. This analysis is performed and presented in Fig. B.29 where the x- and y-axes are the 2D listening location coordinates with color representing the “averaged” MAE for each location.

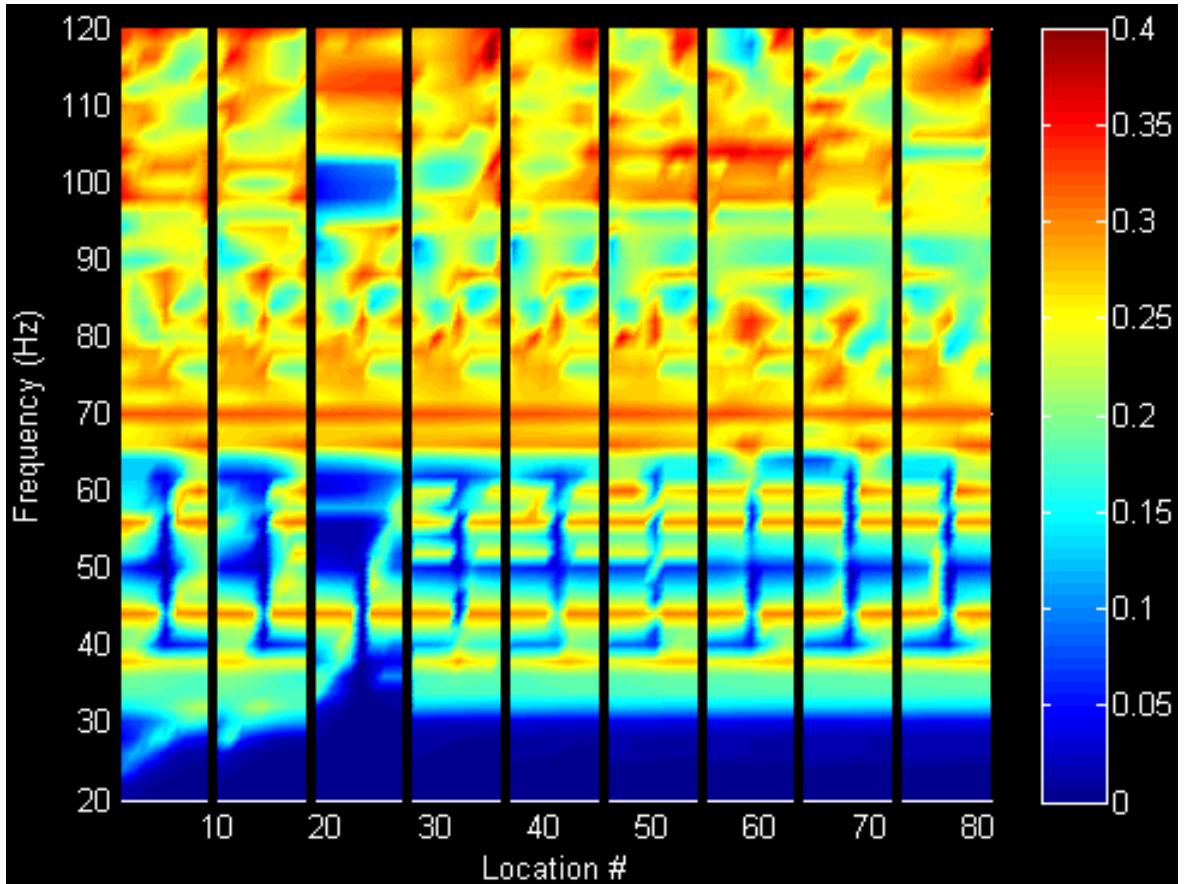


Fig. B.28 Example of an MAE vs. listening location vs. frequency plot for tone burst sweep analysis (81 listening locations, nine rows of nine)

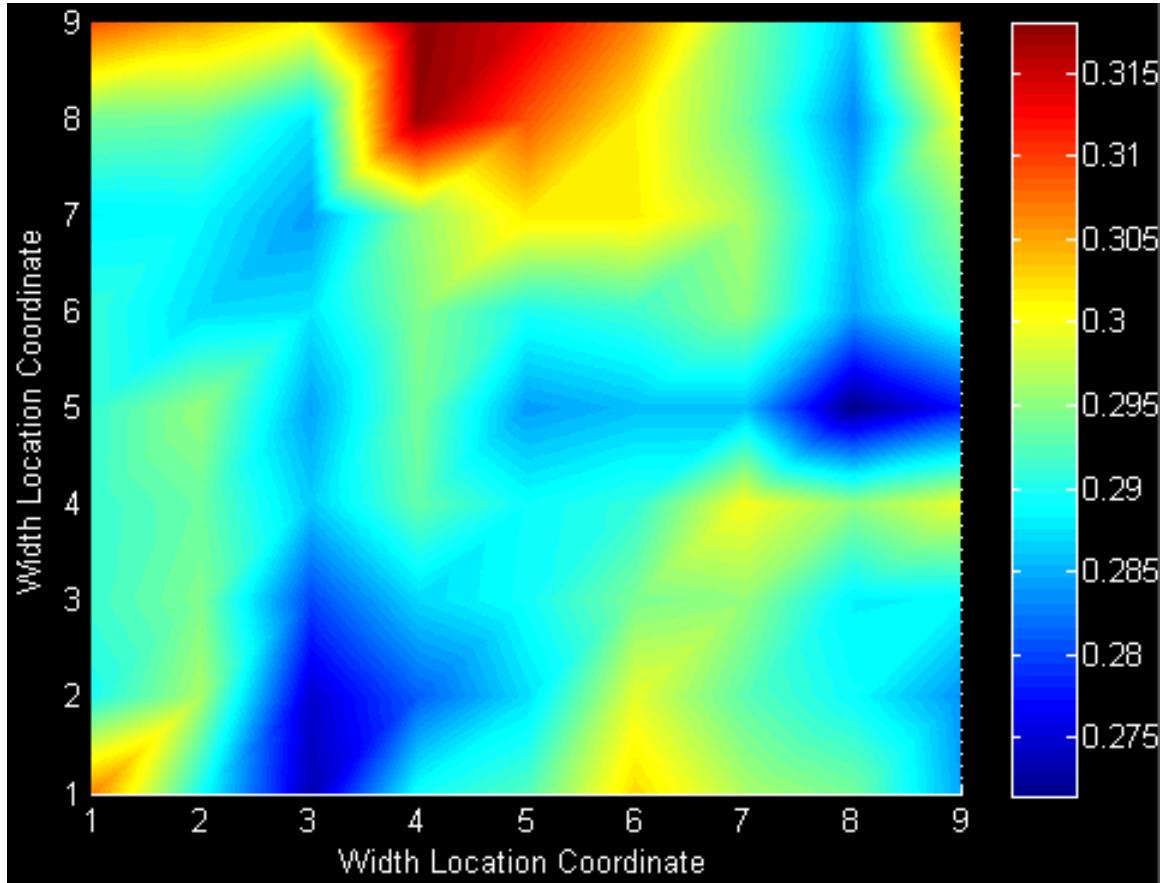


Fig. B.29 Example of average MAE vs. listening location tone burst sweep analysis (listening locations given in terms of listening grid coordinates)

Overall, the presented tone burst analysis functionality of the toolbox provides an additional dimension to room response analysis. In addition to the standard steady-state analysis options commonly used in research, the tone burst data analysis allows direction inspection of the transient response of a sound reproduction system in a non-anechoic space. The analysis is not unlike steady-state spatial variance analysis, where data plots indicate frequencies and/or areas that suffer from great variation over a listening area. The plotting capabilities from both domains should be used side-by-side to develop the best possible low-frequency room-mode correction solution.

B.4 Worked examples

The FDTD toolbox provides extremely high flexibility in terms of simulation possibilities. A sampling of the possibilities is highlighted in this section, primarily concentrating on the non-rectangular topology capabilities of the toolbox, but applied for a wide-range of applications.

B.4.1. Modal buildup animations

It can be very useful to visually examine the nature of a specific modal buildup within a space. This gives clear indication as to which surfaces a standing wave exists between, possibly highlighting areas where some form of bass trap or absorption may be useful or, alternatively, demonstrating the need to adjust the positioning of one or more subwoofer. An example progression of modal buildup is shown in Fig. B.30.

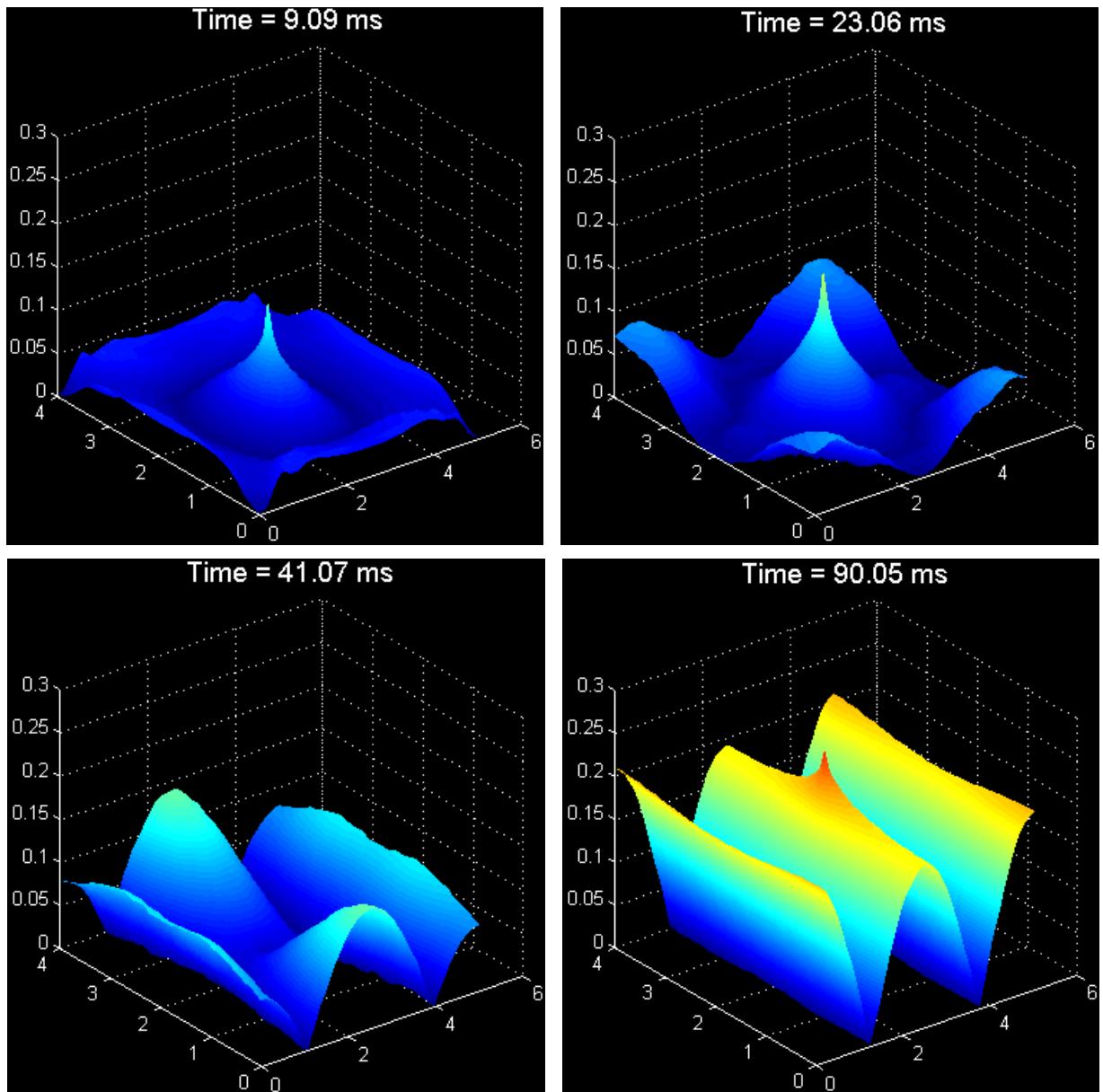


Fig. B.30 Steady-state buildup of an axial mode at 68.8 Hz in a 2D 5 m x 4 m room

B.4.2. Non-rectangular topologies

A key feature of the FDTD toolbox is the non-rectangular topology capabilities. As discussed earlier in this appendix, the default rectangular space can be manipulated using a “cookie-cutter” approach to achieve virtually any variety of topology and/or to include room obstacles. Four examples are included in Fig. B.31 below, highlighting this functionality.

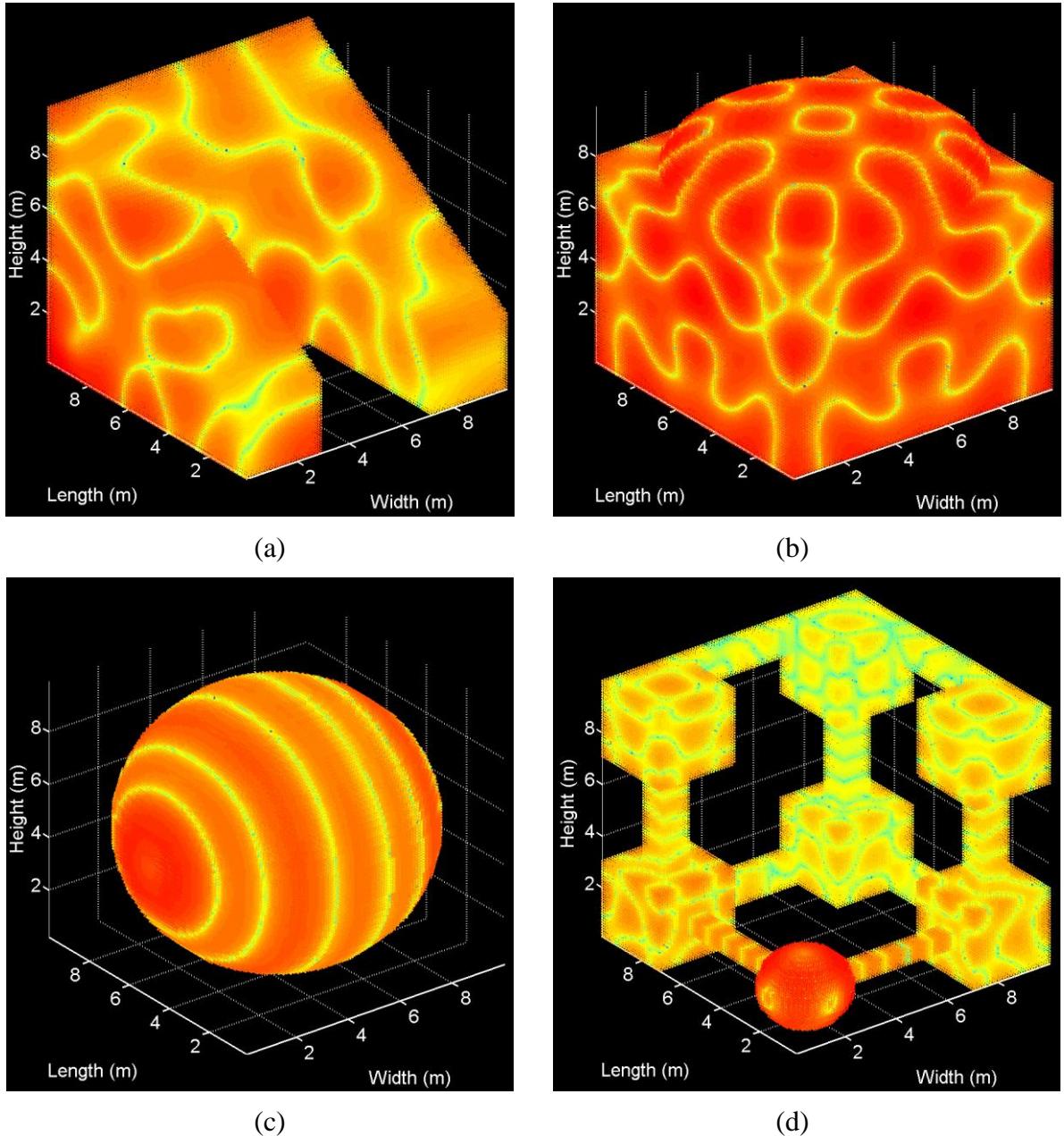


Fig. B.31 Examples of non-rectangular room simulations with a single source driven at (a) 100 Hz located at (1.0 m, 9.0 m, 1.0 m), (b) 120 Hz located at (5.0 m, 5.0 m, 1.0 m), (c) 100 Hz located at (1.0 m, 5.0 m, 5.0 m) and (d) 200 Hz located at (1.5 m, 1.5 m, 1.5 m)

B.4.3. Network of rooms

Along the same lines as the non-rectangular topology examples in the previous section, the toolbox allows the modeling of a network of rooms. This can be used for a range of explorations including Helmholtz resonances due to coupled spaces, the effect of opening a door or window in a room or even the effects of an air duct.

Olson explores the effects of room coupling based on the size of the opening joining two spaces [10]. This well-known research on diffraction can be used as further validation to the FDTD toolbox by simulating similar configurations (Figs. B.32 & B.33).

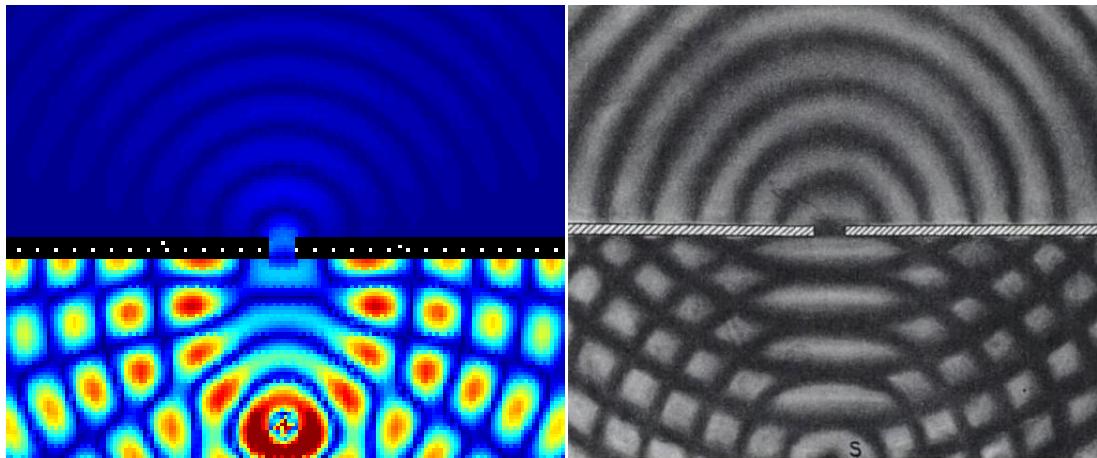


Fig. B.32 Sound diffraction example with a small opening in a reflecting wall
(simulation = left, Olson [10] = right)

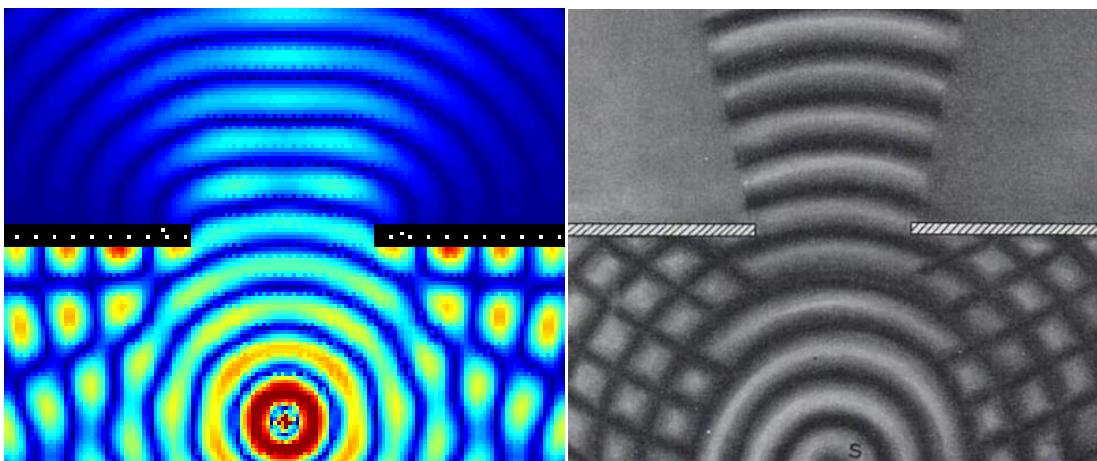


Fig. B.33 Sound diffraction example with a large opening in a reflecting wall
(simulation = left, Olson [10] = right)

Alternatively, this modeling capability can be utilized to examine the complicated interactions between many interconnected rooms. This can be of use if an extremely thorough analysis is necessary for a multiple-room sound reproduction system to avoid any unwanted coupling which could introduce additional modal behavior (Fig. B.34).

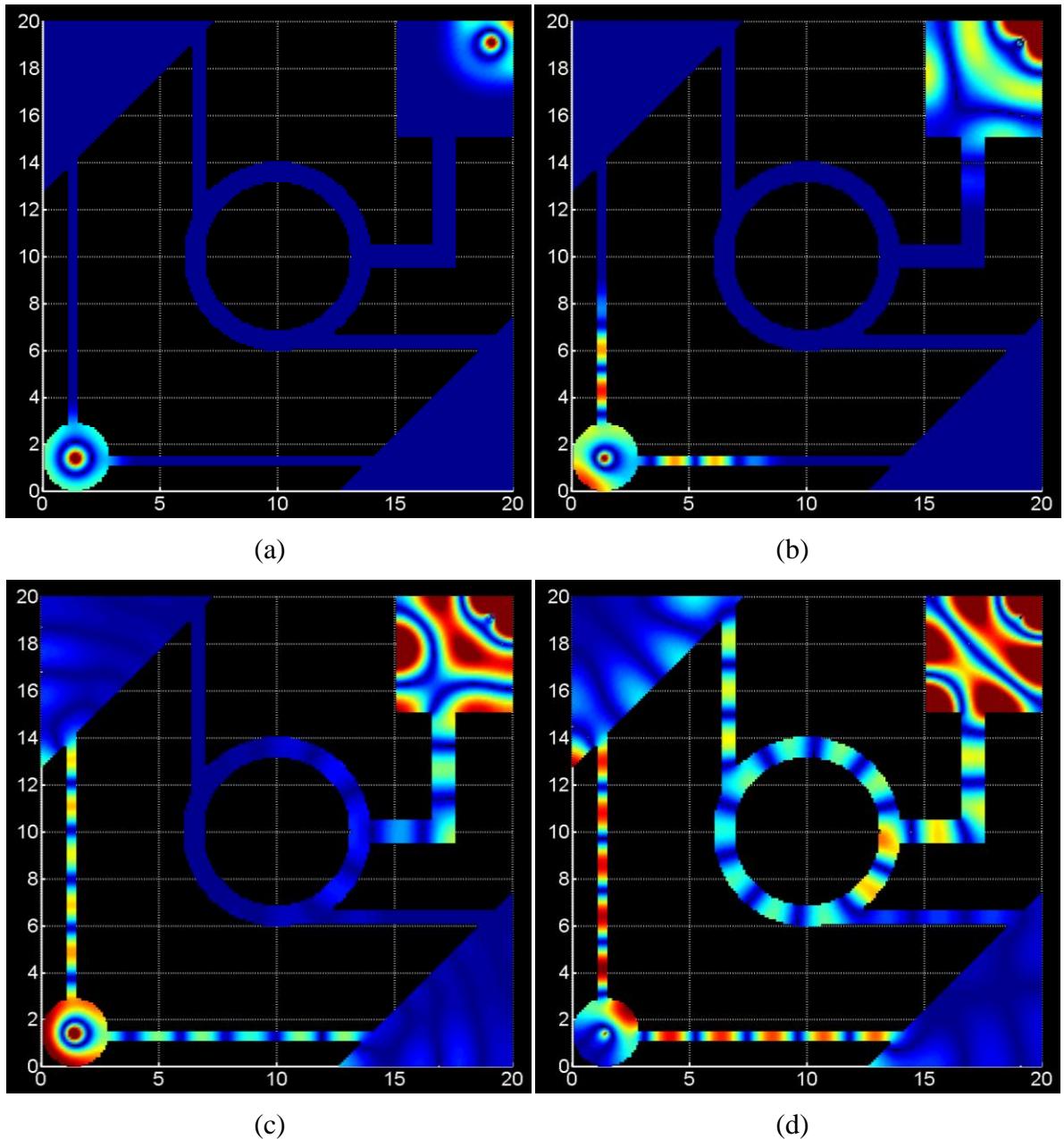


Fig. B.34 Animation snapshots of the simulation of a network of rooms at (a) 10.0 ms, (b) 25.5 ms, (c) 64.0 ms and (d) 2000 ms with two sources at (1.5 m, 1.5 m) and (19.0 m, 19.0 m) driven by an constant 80 Hz sinusoidal signal

B.4.4. Subwoofer polar pattern control visualizations

As discussed in Chapter 4.2.4, multiple drive-units within a single enclosure grant the ability to generate directional low-frequency polar patterns. This capability is built into the FDTD toolbox in the ‘Source Settings’ section with the ‘Polar Pattern’ box. The available patterns are: omnidirectional (default), dipole and cardioid. The dipole and cardioid patterns can be rotated in 30 degree increments, as needed. Simulated examples of each of the available patterns are displayed in Fig. B.35.

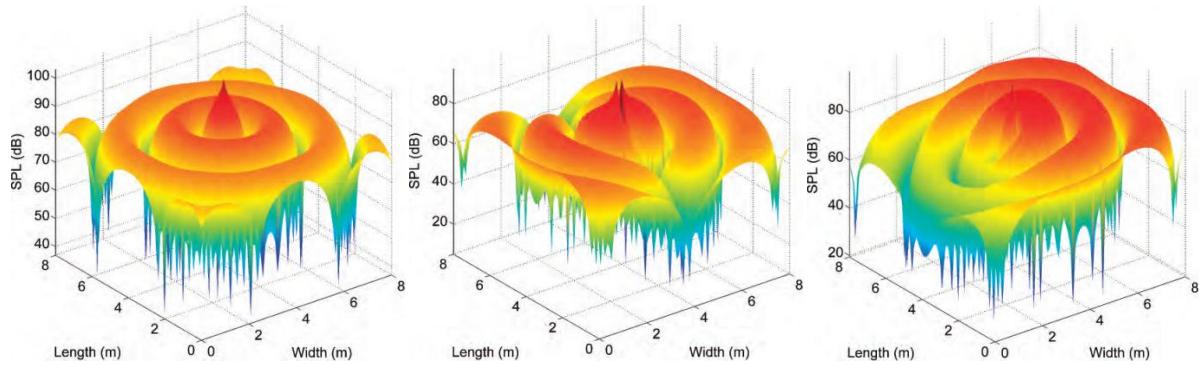


Fig. B.35 Examples of polar pattern control including omnidirectional (left), dipolar (center) and cardioid (right) pattern simulations

B.4.5. Loudspeaker enclosure modeling

An additional analysis option examines the diffraction caused by loudspeaker enclosures. This analysis feature is accomplished by inserting obstacles within a room that intersect to create a box. A source is then positioned at a small opening on one side to approximate a drive-unit. Cabinet diffraction is relevant over all frequencies since the geometry of the cabinet remains unchanged; however, its manifestation differs with frequency. At frequencies with wavelengths shorter than the average surface radius of the cabinet, diffraction causes lobing, thus making critical the listening angle.

A simple example is given in Fig. B.36 with an 0.8 m by 0.8 m sealed enclosure driven at 250 and 800 Hz. It is possible to examine a nonrectangular enclosure configuration, such as an arbitrary elliptical design (Fig. B.37). This can help examine whether a special enclosure design will reduce diffraction, giving a more even polar pattern. The example highlighted in Fig. B.37 shows a decrease in diffraction at both 250 and 800 Hz due to the rounded shape of the enclosure, reducing the diffraction effects caused by sharp cabinet edges.

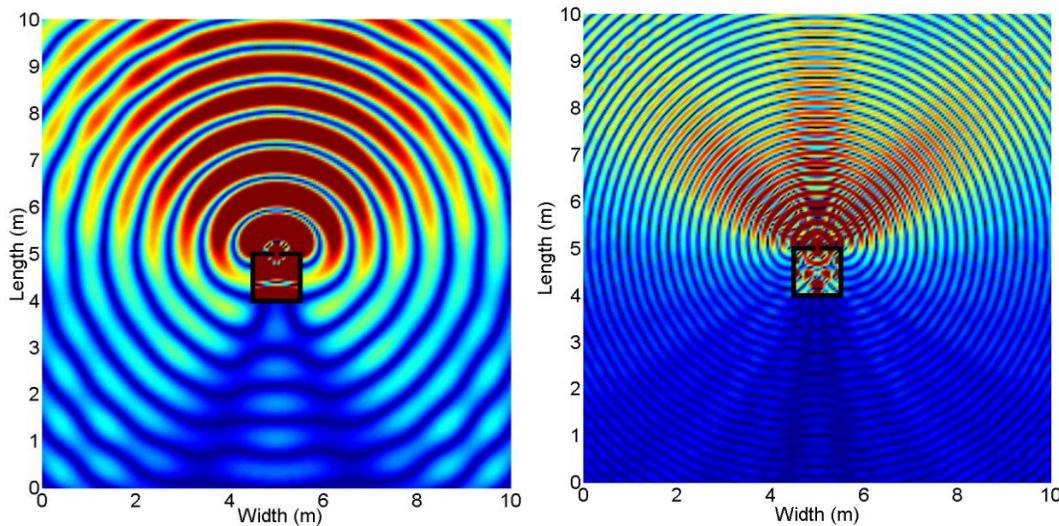


Fig. B.36 Simulation of a rectangular sealed enclosure at 250 Hz (left) and 800 Hz (right)

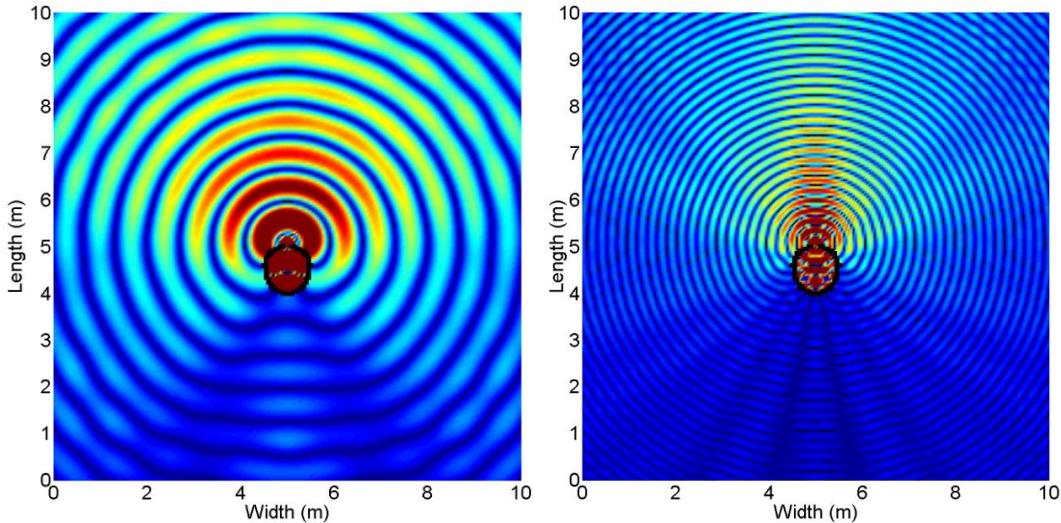


Fig. B.37 Simulation of an elliptical enclosure at 250 Hz (left) and 800 Hz (right)

The lower frequency range experiences a different effect from diffraction due to the longer wavelengths, resulting in closer similarity in phase of the diffracted field to the direct sound. In this frequency band diffraction will cause a shift in the acoustic center of the loudspeaker. This concept has been explored by Vanderkooy [21, 22], showing that below 200 Hz cabinet diffraction causes the acoustic center to be located a distance in front of the drive unit, relative to the front surface width and cabinet depth. An example of this effect is given in Fig. B.38, which shows a 0.5 by 0.5 m enclosure driven at 50 Hz.

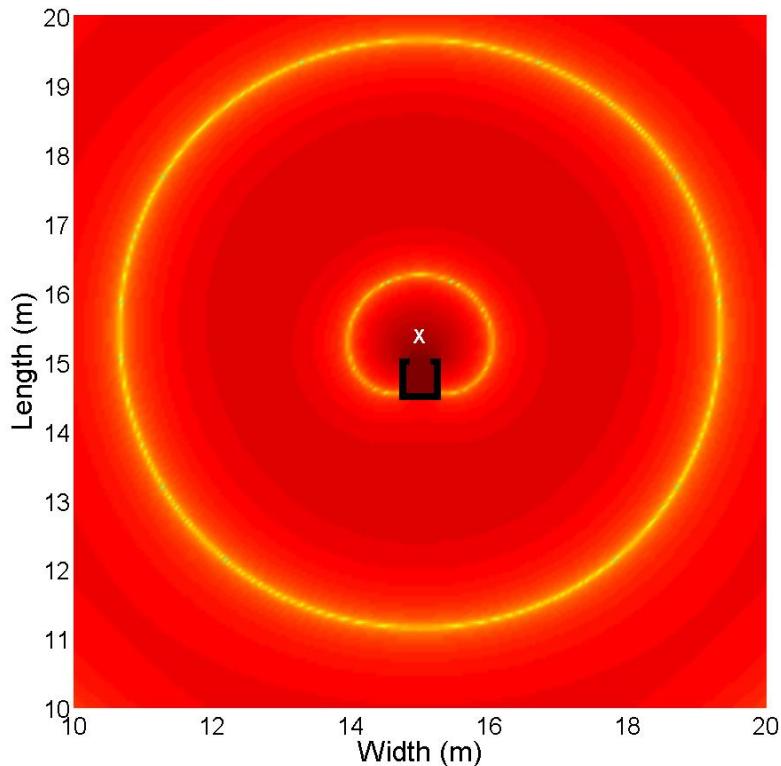


Fig. B.38 Simulation of a 50 cm x 50 cm rectangular enclosure at 50 Hz showing a forward acoustical center shift (indicated by the X)

Appendix C – Journal publications

This section contains a list of the three journal publications related to the research presented in this thesis. High-resolution color versions of these publications are included on the CD-ROM attached to the rear page of the thesis.

- 1) Hill, A.J.; M.O.J. Hawksford. “Visualization and analysis tools for low-frequency propagation in a generalized 3D acoustic space.” *Journal of the Audio Engineering Society*, vol. 59, no. 5, pp. 321-337. May, 2011.
- 2) Hill, A.J.; M.O.J. Hawksford. “Chameleon subwoofer arrays in live sound.” *Institute of Acoustics, Acoustics Bulletin*, vol. 36, no. 3, pp. 32-40. May/June, 2011.
- 3) Hill, A.J.; M.O.J. Hawksford. “Wide-area psychoacoustic correction for problematic room-modes using nonlinear bass synthesis.” *Journal of the Audio Engineering Society*. Scheduled publication: November, 2011.

Appendix D – Conference papers

This section contains a list of presented conference papers relating to the topics discussed in this thesis. High-resolution color versions of papers are included on the CD-ROM attached to the rear page of the thesis.

- 1) Hill, A.J.; M.O.J. Hawksford; A.P. Rosenthal; G. Gand. "Subwoofer positioning, orientation and calibration for large-scale sound reinforcement." 128th Convention of the Audio Engineering Society, London, UK. May, 2010.
- 2) Hill, A.J.; M.O.J. Hawksford. "A hybrid virtual bass system for optimized steady-state and transient performance." Proc. 2nd Computer Science and Electronics Engineering Conference (CEEC'10), IEEE Xplore, Colchester, UK. September, 2011.
- 3) Hill, A.J.; M.O.J. Hawksford. "Individualized low-frequency response manipulation for multiple listeners using chameleon subwoofers arrays." Proc. 3rd Computer Science and Electronics Engineering Conference (CEEC'11), IEEE Xplore, Colchester, UK. July, 2011.
- 4) Hill, A.J.; M.O.J. Hawksford. "Visualization and analysis tools for low frequency propagation in a generalized 3D acoustic space." 127th Convention of the Audio Engineering Society, New York, NY, USA. October, 2009.
- 5) Hill, A.J.; M.O.J. Hawksford. "Seeing through the darkness: Visualizing low frequency behavior in small rooms." Proc. Institute of Acoustics – Conference on Reproduced Sound, Cardiff, UK. November, 2009.
- 6) Hill, A.J.; M.O.J. Hawksford. "Chameleon subwoofer arrays – Generalized theory of vectored sources in a closed acoustic space." 128th Convention of the Audio Engineering Society, London, UK. May, 2010.
- 7) Hill, A.J.; M.O.J. Hawksford. "Multi-band low-frequency room correction with chameleon subwoofer arrays." Proc. European Acoustics Association – EUROREGIO 2010, Ljubljana, Slovenia, September, 2010.
- 8) Hill, A.J.; M.O.J. Hawksford. "Wide-area psychoacoustic correction for problematic room modes using non-linear bass synthesis." 129th Convention of the Audio Engineering Society, San Francisco, CA, USA. November, 2010.
- 9) Hill, A.J.; M.O.J. Hawksford. "Chameleon subwoofer arrays in live sound." Proc. Institute of Acoustics – Conference on Reproduced Sound, Cardiff, UK. November, 2010.
- 10) Hill, A.J.; M.O.J. Hawksford; A.P. Rosenthal; G. Gand. "Kick drum signal acquisition, isolation and reinforcement optimization in live sound." 130th Convention of the Audio Engineering Society, London, UK. May, 2011.