Validation of a 3-D Virtual Acoustic Prototyping Method For Use In Structural Design

Zachary T. Carwile

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
In partial fulfillment of the requirements for the degree of

Master of Science
In
Mechanical Engineering

Dr. Marty Johnson, Chair
Dr. Jamie Carneal
Dr. Mark Paul

January 17, 2006
Blacksburg, Virginia

Keywords: Virtual acoustic prototyping, Structural acoustics,
Equivalent source method, HRTF

Validation of a 3-D Virtual Acoustic Prototyping Method For Use In Structural Design

Zachary T. Carwile
Mechanical Engineering Department
Virginia Tech

Abstract

Virtual acoustic prototyping (auralization) is the rendering of a virtual sound field that is created from the calculated acoustic response of a modeled structure. Auralization is useful in the design and subjective evaluation of buildings, automobiles, and aircraft. The virtual acoustic prototyping method in this thesis uses finite element modeling (FEM), the equivalent source method (ESM), and head-related transfer functions (HRTFs). A tradeoff exists between the accuracy of the auralization process and the number of equivalent sources (and thus computational power) that are required.

The goal of this research is to validate (numerically and subjectively) a virtual acoustic prototyping method for use in structural design; this thesis illustrates the first attempt to apply the aforementioned methods to a structure that represents a typical building or automobile. The structure's acoustics were modeled using FEM, ESM, and HRTFs. A prototype of the modeled structure was built. A 36% correlation was achieved between the model and prototype. Slight variations in boundary conditions caused significant FEM error, but the data represented a typical structure.

Psychoacoustic comparison testing was performed to determine the number of equivalent sources that must be used in an auralization to accurately recreate the sound field. The number was found to be dependent on the type of noise that is played to the test subject. A clear relationship between the numerical correlation of two sounds and the percentage of subjects who could hear a difference between those two sounds was established for impulsive, broadband, and engine noises.

Acknowledgements

I would like to thank my advisor, Dr. Marty Johnson, for hiring me to complete this project, teaching me about acoustics, and overseeing my work. In addition, I am grateful for the direction and support this project has received from AuSIM, Inc. in Mountain View, California. I would also like to thank the staff and student researchers of the Virginia Tech Vibration and Acoustics Laboratories (VAL), particularly Dr. Jamie Carneal for assistance with equipment and technical matters; Dan Mennitt for the preparation of some data and figures in Section 2.3; and Melissa Williams for her support of myself a the rest of VAL. A thank you also goes out to the all students and faculty of VAL for performing my psychoacoustic testing and generally supporting my project. Finally, I would like to thank friends I have made in my 5-½ years at Virginia Tech and my family for supporting me in this educational endeavor.

Table of Contents

List of Multimedia Objects

List of Tables

List of Figures

Nomenclature

a_0	linear-fit coefficient, constant term
a_1	linear-fit coefficient, first-order term
A	cross-sectional area
b_L	time-domain binaural signal at the left ear
b_R	time-domain binaural signal at the right ear
b_i	arbitrary time-domain binaural signal
B_L	frequency-domain equivalent of left binaural signal
B_R	frequency-domain equivalent of right binaural signal
c	speed of sound
C	damping matrix
D	occurrence of perceived differences (percentage)
e	exponential function
e_{LQSR}	least-squares algorithm error
E	modulus of elasticity
$EPWL$	evaluation points per acoustic wavelength
ESM	equivalent source method
f	frequency (Hertz)
$\mathbf{F}$	force vector
FEM	finite element modeling
H_L	Left head-related transfer function
H_R	Right head-related transfer function
$HRIR_L$	Left head-related impulse response
$HRIR_R$	Right head-related impulse response
HRTF	head-related transfer function
i	index
j	square-root of -1
k	wave number (except for Section 3.1)
k, k_i	element stiffness (Section 3.1)

K	stiffness matrix
KEMAR	Knowles Electronic Manikin for Acoustics Research
L, L_i	beam length
LSQR	least-squares algorithm
m, m_i	element mass (Section 3.1)
m	mass (Section 3.2)
M	mass matrix
N	numerical correlation between two time signals
NES	number of equivalent sources
p	acoustic pressure
P	axial beam loading (Section 3.1)
P	atmospheric pressure (Section 3.2)
$P_i(\omega)$	nodal pressure from finite element model
P	pressure vector at evaluation points
q_i	time-domain equivalent source strength
Q	volume velocity
$Q_i(\omega)$	frequency-domain equivalent source strength
Q	equivalent source strength vector
r	radius of an acoustic radiator
r_e	radius of the evaluation sphere
R	gas constant for air (Section 3.2)
R, R_{ij}	source-to-receiver distance (except for Section 3.2 and Chapter 5)
R	correlation coefficient (Chapter 5)
S	surface area of an acoustic radiator
$SPWL$	equivalent sources per acoustic wavelength
t	time
t_0	time delay
T	temperature (Section 3.2)
T, T(ω)	transfer function between **P** and **Q** (except for Section 3.2)
u	surface velocity of an acoustic radiator
x	displacement vector

x_i	element displacement
$\dot{\mathbf{x}}$	velocity vector
$\ddot{\mathbf{x}}$	acceleration vector
$\ddot{x}_i$	element acceleration
V	volume
z	number of points in the evaluation volume
α	mass-proportional damping coefficient
β	stiffness-proportional damping coefficient
δ	beam deflection
ϕ_H	head elevation angle
ϕ, ϕ_i	elevation angle
λ	wavelength
μ	mean
θ_H	head azimuth angle
θ, θ_i	azimuth angle
ρ	density
ρ_0	density of air
σ	standard deviation
υ	Poisson's ratio
ω	frequency (radians)
ω_i	natural frequency (radians)
ζ, ζ_i	damping ratio

Chapter 1: Introduction

The objective of this thesis is to summarize the author's research in the field of virtual acoustics. The work performed relates specifically to the validation of a virtual acoustic prototyping method for use in structural design. In order to understand the motivation and scope of this research and the current literature on this topic, a brief introduction to virtual acoustics is given.

1.1 Introduction to Virtual Acoustics

Virtual acoustics is the study and practice of creating virtual sound fields; it is part of the broader practice of virtual reality, which is the process of using a computer to create a simulated environment. Virtual reality usually consists of two sensory components – visualization and auralization – and the latter uses virtual acoustics. Kleiner defines auralization as "the process of rendering audible, by physical or mathematical modeling, the sound field of a source in a space, in such a way as to simulate the binaural listening experience at a given position in the modeled space." [1] Binaural sound contains directional cues that are intended to give the perception that the sound emanates from outside of the true source. A subject can listen to a simulation using a set of headphones and hear audio cues that appear to be coming from specific directions and from far outside of the head.

Virtual acoustic prototyping is the application of virtual acoustics as a design tool during the development phase of a structure. Noise levels can be predicted virtually without constructing a physical prototype, which can greatly reduce the cost of developing a new product. Since noise and sound quality are complex and subjective measures, effective auralization of a virtual acoustic prototype is a challenging task.

A flowchart of the virtual acoustic prototyping method used in this thesis is shown in Figure 1.1. The process consists of three components: structural-acoustic modeling, the equivalent source method (ESM), and the application of head-related transfer functions (HRTFs). The acoustics of the structure are modeled using one of many

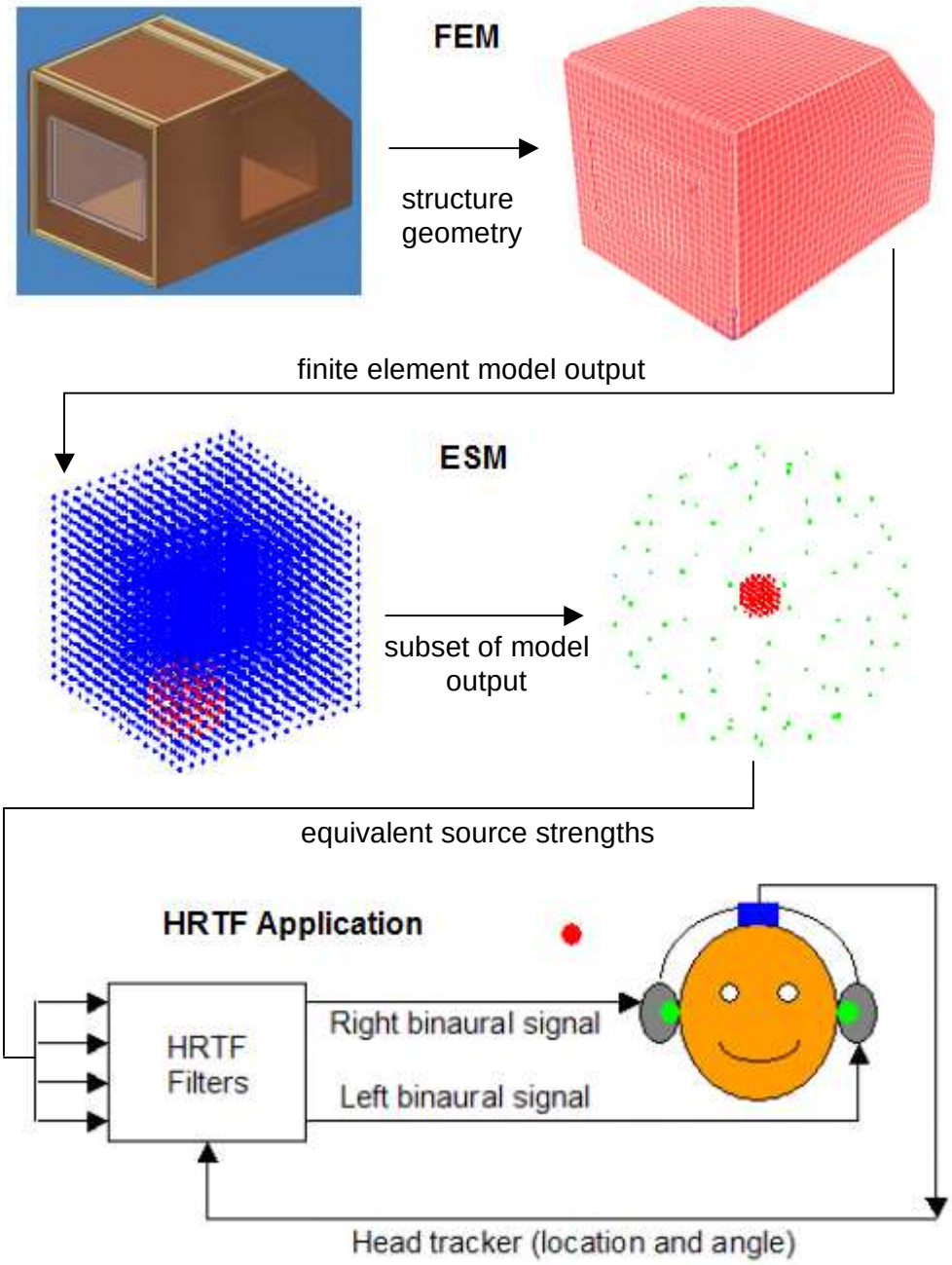

Figure 1.1. Flowchart of the virtual acoustic prototyping method used in this thesis.

available techniques, such as the finite element method (FEM), the boundary element method (BEM), ray tracing, or statistical energy analysis (SEA). The information can also be obtained from an analytical model or measured data; the acoustic data must be complex (containing both magnitude and phase information). FEM is the primary modeling technique used in this thesis, so it will be the only method discussed in detail.

The geometry and material properties of the acoustic space and surrounding structure are specified in a dedicated FEM software package. The loads (i.e. force on structure, monopole source) and boundary conditions are then specified. A meshing tool is used to divide each part into hundreds or thousands of finite elements. Finally, the response of the acoustic space is determined at a specified set of frequencies. In acoustic modeling, the characteristic length of an element should be less than one sixth of the acoustic wavelength at the highest frequency in question. Because of this general rule, finite element modeling of large structures is only effective at low frequencies. [2] Figure 1.2 shows an example of the output for a finite element model of an acoustic space, surrounded by ¾"-thick medium-density fiberboard (MDF), that is acted on by a Lexan plate. The acoustic pressure field is solved for at a set of frequencies; the total acoustic response of the space is achieved by summing the pressure fields for each particular frequency.

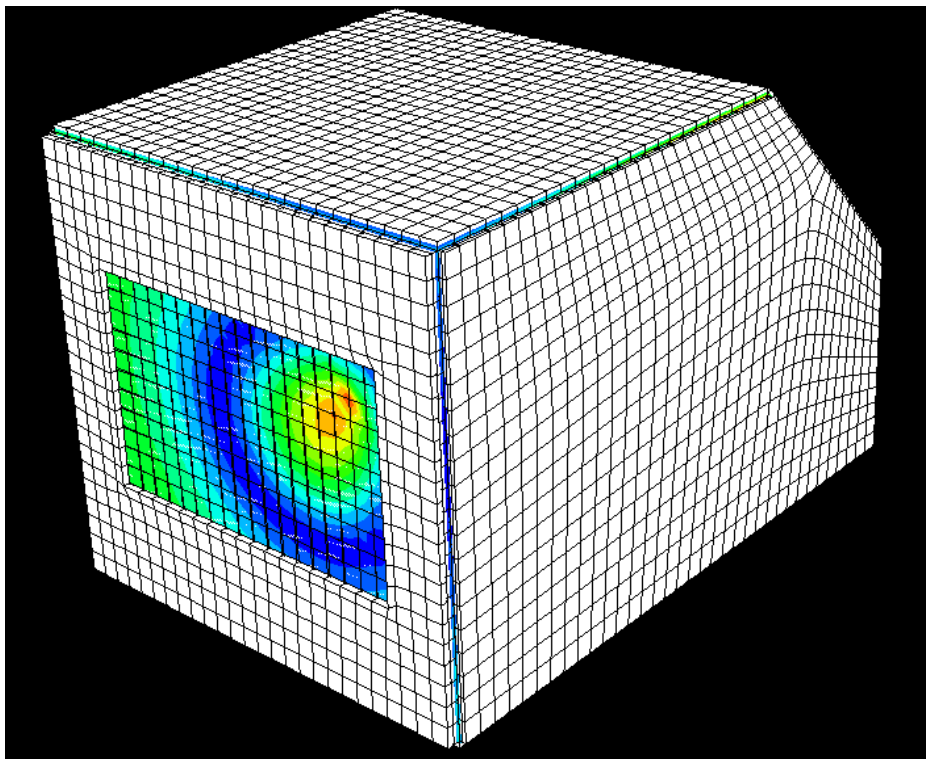

Figure 1.2. Finite element model output. An acoustic medium is surrounded by MDF (white) and a vibrating Lexan window (clear). The acoustic pressure response is shown; red indicates high pressure and blue indicates low pressure.

In order to produce binaural sound, the direction from which the sound is coming needs to be known. In a structural-acoustic system, every point on every moving surface acts as a source of sound. If the structure is simple, such as a vibrating plate, the surface can be discretized and each patch can be treated as an equivalent source. [3] A real structure, such as an automobile, is complicated by large surfaces, reverberation, and absorptive material. Treating each surface and reverberation as many equivalent sources would be computationally prohibitive. Therefore, the complex acoustic response is calculated using FEM or another method, and that response is recreated using a much smaller number of sources in the equivalent source method.

The finite element data is in the form (x, y, z, p), where x, y, and z represent spatial coordinates and p represents the acoustic pressure at that nodes of the mesh; no directional information is contained in the data. ESM is used to give directional information to the sound field. A subset of the finite element data, termed the evaluation volume, is recreated using virtual monopole sources, or equivalent sources. The sources are placed on a large sphere around the evaluation as shown in Figure 1.3. An iterative solver is used to determine the amplitude of each equivalent source that is necessary to recreate the sound field (x, y, z, p). Large numbers of nodes require many equivalent sources, so often only a small portion of the acoustic pressure field is recreated using ESM.

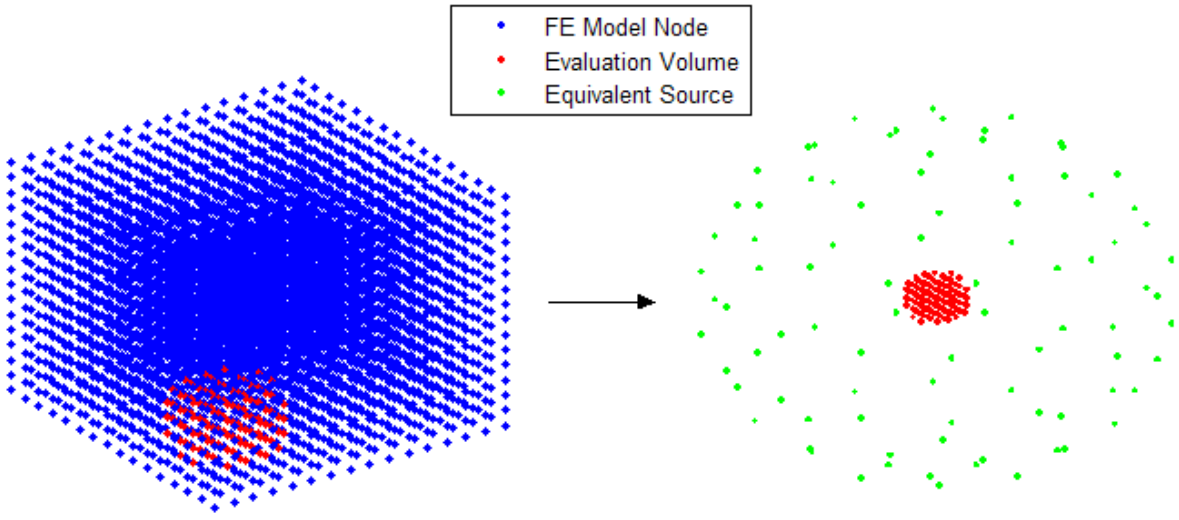

Figure 1.3. ESM attempts to recreate a pressure field (red), whose source may be a vibrating structure, using simple monopole sources (green) distributed on a sphere.

To complete the auralization, HRTFs are applied to the equivalent sources. HRTFs are measurements of the direction-dependent acoustic filtering that a free-field sound experiences due to the head, torso, and pinnae (outer ears). [4] HRTFs contain information about the interaural time difference (ITD) and interaural intensity difference (IID) between the two ears. HRTFs are a function of ear (left or right), elevation angle, and azimuth angle, and sound frequency. A head-related impulse response (HRIR) is the time-domain equivalent of an HRTF. The HRIR is the acoustic response in the ear canal due to an impulse sound source at a particular direction. Figure 1.4 illustrates the concept of the HRIR; HRTFs will be described in greater detail in Chapter 2.

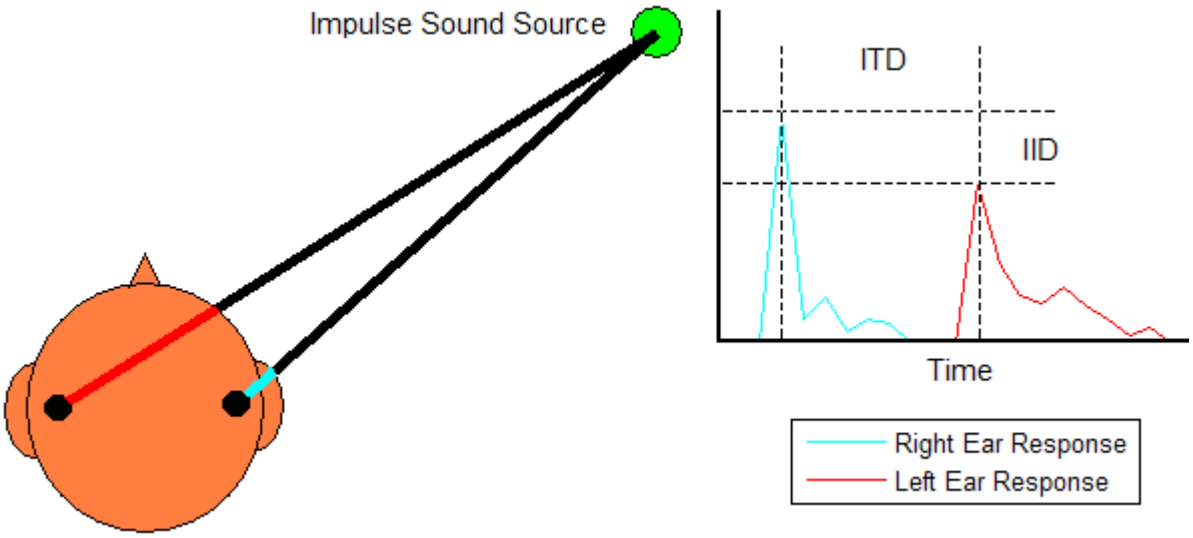

Figure 1.4. Relationship between HRIR, ITD, and IID. The sound first reaches the right ear, and after microseconds, it reaches the left ear with a slightly lower amplitude.

The appropriate HRTFs are applied to each equivalent source; since the equivalent sources may not all be the same distance from the listener, a time delay and gain is applied to each source. The time domain source signals are superimposed and played over a set of headphones. The listener will be able to hear the modeled structure, and the sound source will appear to be coming from the correct location. If a head-tracker is used, the listener can move and rotate and the sound will still appear to be at the same location. Figure 1.5 illustrates this phenomenon. The equivalent sources are constant, but as the head is rotated, the head tracker output triggers a change in the HRTF filters. Thus, as the head moves, the actual sound sources (headphone speakers shown in

green) move with the head, but the apparent sound source (shown in red) does not change.

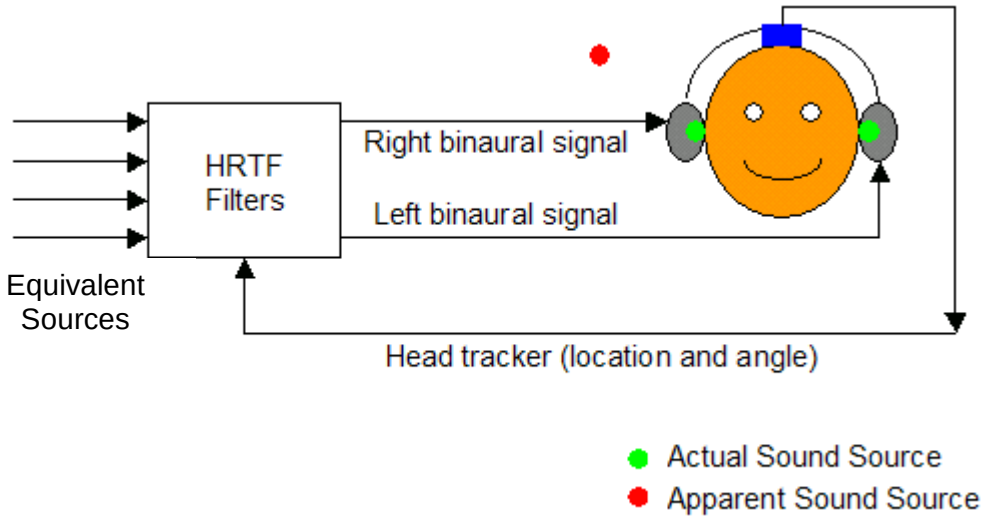

Figure 1.5. Auralization through the use of HRTFs allows for the production of sound that appears to emanate from outside of its true source.

This virtual acoustic prototyping method has only been tested numerically; no subjective tests have been performed. Noise and sound quality are highly biased measures, subjective validation of this method is imperative.

1.2 Motivation for Work

Commercially available acoustic simulation software allows a user to predict natural frequencies, identify vibration hot spots, and solve for the pressure response. Auralization packages are available that let the user model a structure and subsequently experience its sound. Currently, no auralization software is available that allows the user to perform a head-tracked acoustic simulation in real time; this would be a practical tool for virtual acoustic prototyping.

Virtual acoustic prototyping is useful in the design and subjective evaluation of buildings, automobiles, and aircraft. Time and money can be saved by predicting sound problems during the structural design phase rather than fixing problems after construction or mass-production. Since noise, vibration, and harshness have adverse effects on human health, resources can be saved in the long run as well.

In aircraft and automotive design, acoustic engineering has generally taken place after production. Millions of dollars are spent annually to find ways to reduce internal and external noise emanating from aircraft. Hundreds of millions of dollars are spent optimizing the appearance and performance of a new car designs; some are immediately shunned by automotive editors because of excessive cabin noise or a rough engine note [5], while others are hailed for a sporty engine or exhaust note. Car and Driver magazine recently praised the high-performance 2005 Pontiac GTO: "Exhaust gases flow freely, and the drum beat is enough to cause meek dogs to cower." [6]

In building design, sound sources have been inadvertently placed too close to work areas. A mechanical room in Virginia Tech's Durham Hall was placed across the hall from a row of offices; the work environment becomes noisy whenever the elevator is used. In home theatre design, even the best speakers will sound unclear if the acoustic properties of the room are substandard. Virtual acoustic prototyping seeks to predict and eliminate these problems at the minimum cost.

An effective, real-time, head-tracked auralization method would be a welcome enhancement for virtual acoustic prototyping. To perform a head-tracked acoustic simulation, the ESM is required in order to apply HRTFs to the finite element data; the result is a binaural signal. Software has been developed that implements the ESM. The time needed to compute the equivalent source magnitudes can varies depending on the number of sources. The number of equivalent sources can be increased or decreased by altering the density of sources or by changing the size of the evaluation volume. Having more equivalent sources typically produces greater accuracy. The objective of this work is to determine the level of accuracy at which a listener can hear a difference between a sound and its ESM reproduction.

1.3 Review of Literature

The research outlined in this thesis is the subjective validation of a 3-D virtual acoustic prototyping method. The purpose of a virtual acoustic prototyping package is to allow a user to experience the acoustics a structure during the design phase. Extensive

research has been performed in the areas of acoustic modeling and the application of HRTFs to produce 3-D sound. However, the research on the auralization of acoustic models is more limited. Detailed studies concerned with the subjective validation of virtual acoustic prototyping methods have not been performed.

Currently, the most available auralization system is fully-computed auralization (FCA). [1] Figure 1.6 is a diagram of a FCA system. FCA has not been applied to a modeled-structure in a real-time head-tracked environment, but with proper development, this is an achievable goal. The structure in question will be modeled on a computer to calculate the binaural room impulse response (BRIR), which is the acoustic response of the room to an impulse source. The source can be either sound or a force on the structure. A digital signal processing algorithm will be used to convolve the BRIR with an input signal. Auralization is achieved by presenting the resulting binaural signal to a listener using a set of headphones. Head-tracking can be used to enhance an auralization; the sound will appear to be coming from the same location in space regardless of the head's orientation. Therefore, the apparent source of the sound will not move with the head.

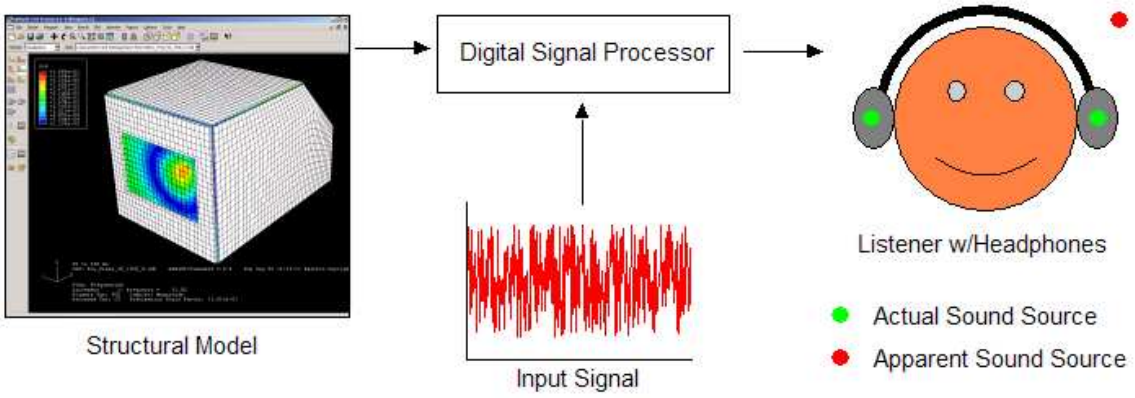

Figure 1.6. Diagram of a fully-computed auralization system.

FCA has been applied at the NASA-Ames Research Center; researchers have developed the Virtual Environment Workstation (VIEW) – an acoustic display that is capable of generating directional sound cues over headphones in real time. [7] In order to localize the sound cues, VIEW uses measured HRTFs rather than computed BRIRs. The

perceptual accuracy of the application of HRTFs in a FCA system has been validated subjectively. However, VIEW cannot be used as a virtual acoustic prototyping system because the sound cues corresponding to structural excitation are not known.

Johnson [8] proposed the equivalent source method (ESM) for determining the sound cues that simulate structural excitation. The ESM uses the principle of wave superposition to recreate a complex acoustic field. [9] The modeled acoustics of a control volume within a structure would be reproduced by a set of virtual monopole sources with a known location and frequency content. These virtual sources – which will from now on be - termed "equivalent sources" – are sound cues necessary for auralization. The application of equivalent source method is similar to the production of virtual orchestras [10], whereby several strategically-placed loudspeakers are used to simulate the acoustics of live orchestra consisting of a much larger number of musicians. The ESM approximates the acoustics of a complex vibrating structure using a relatively small number of monopole sources.

The ESM requires a model of the acoustics of the structure that is auralized. Acoustic modeling of structures has been comprehensively researched. Finite element modeling is the most popular method, which was first applied to structural acoustics problems by Zimmerman and Gladwell [11]. Variational methods were used to formulate the equations governing the harmonic vibration of coupled structural-acoustic systems. Zimmerman [12] applied his method to the simple cases of air-membrane and air-plate coupling.

As computer technology progressed, the finite element method was applied to more complex acoustics problems, and the results were compared to measured data. Petyt and Koopmann [13] modeled rectangular and irregular enclosures and Craggs [14] investigated the sound transmission between two enclosures coupled by a plate. Predicted and measured natural frequencies typically differed by less than 2%. Initially, the finite element method was used only to determine the natural frequencies of a structure. Later, damping was added to models in the form of absorptive elements [15] and wall impedances [16]. Nefske [17] applied finite element techniques to the automotive design process; resonances and panel-to-air transfer functions were calculated

for a particular car model in the frequency range of 20-200 Hz. The car design was altered, reducing noise levels by up to 5 dB.

For frequencies higher than 500 Hz, the finite element method is not desirable. Since wavelengths are smaller, more elements must be used; the model solution is also very sensitive to boundary conditions and material properties. Ray tracing and statistical energy analysis (SEA) are typically more appropriate for high-frequency modeling of structures. In a complex structure, finite element techniques may be appropriate for one part while SEA techniques are appropriate for another part. Mace and Shorter [18] recently developed a hybrid method that can combine finite element data and SEA data. Hopkins [19] and Shorter [20] validated the method with structural models and experimentation. Other methods of structural modeling include the evaluation of scaled models produced by stereo lithography. [21] Although it produces accurate results, this practice is currently limited to determination of natural frequencies; it is suitable for structures of extremely complex geometry. SEA, the hybrid technique, and scale modeling have not been used in conjunction with the ESM. These techniques should be a topic of future research, since they may hold the key to auralization at higher frequency ranges where the finite element method is no longer feasible.

The finite element method, as formulated by Zimmerman, is a paradigm for computing the acoustic pressure field of a structure. The pressure field alone contains no information about the acoustic source location, so the ESM is required to calculate the sound cues that correspond to structural excitation. Equivalent sources are distributed about the surface of a virtual sphere that encircles the pressure field location, and their amplitudes are adjusted to match the pressure field. The number of equivalent sources required is usually much lower than the number of nodes in the finite element model of the pressure field.

The mathematical robustness of the ESM has been validated [22,23], and the error has been quantified extensively using finite element data. [24] However, the method has not been tested subjectively. Psychoacoustics – the study of the human perception of sound – is a well-documented field. [25,26] However, previous research concerned with the human factors involved in the auralization of structural acoustics is limited to interior and exterior aircraft sound. Sullivan et. al. [27] have developed an aircraft interior noise

simulator for psychoacoustic testing at the NASA-Langley Research Center, but auralization is not performed with real-time head tracking. The accuracy of the auralization method depends on the number of equivalent sources; there is a tradeoff between computation time and mathematical error. It is possible that there is a diminishing return for increasing the number of equivalent sources – continuing to increase the number of sources may not change the listener's perception of the simulation. The relationship between the mathematical accuracy of the ESM and the perceptual accuracy of the binaural signal is currently unknown.

Acoustic modeling, the ESM, and the application of HRTFs have been researched almost exhaustively, the research in this thesis describes the mathematical validation of a complete virtual acoustic prototyping system. In addition, the psychoacoustic robustness of the ESM and the subsequent application of HRTFs is presented.

1.4 Scope and Relevance of Work

The goal of this research is to validate – both numerically and subjectively –a virtual acoustic prototyping method for use in structural design. This thesis illustrates the first attempt to apply the aforementioned methods to a structure that represents a typical building or automobile.

In the first part of this research, two models were created – a physical model and a numerical model. For the numerical model, FEM, ESM, and HRTFs were applied to a hypothetical structure. For the physical model, the hypothetical structure was actually built. A manikin was placed inside and the acoustic response at the ears was measured. The outputs of the numerical and physical models are compared.

In the second part of this research, subjective tests were performed to determine what ESM error level is audible to the average listener. The exact reproduction of a complex sound field using ESM is not feasible; the error level depends primarily on the number of equivalent sources that are used in the reproduction. Computing power dictates the upper limit on the number of equivalent sources that can be used in an auralization. The addition of more sources reduces the numerical error, but the human

ear may not notice the small reduction in error. The goal of the subjective testing was the determine the point of diminishing return in ESM.

1.5 Outline of Thesis

Chapter 2 of this thesis will expand on the introduction to virtual acoustics in Section 1.1; HRTFs, ESM, and the numerical error associated with the virtual acoustic prototyping process are explained in greater detail. Chapter 3 describes the numerical modeling of virtual acoustic prototyping for a particular structural acoustic test platform (which is frequently referred to as "structure" or "enclosure"). The structure is modeled using FEM and, subsequently, ESM is applied to an evaluation sphere positioned in the enclosed acoustic space. HRTFs are applied to a hypothetical manikin residing inside the sphere. Chapter 4 describes the numerical validation of the model from Chapter 3. A real structure is built, and a manikin is inserted; the acoustic response at the ears is measured and compared to the numerically-modeled response. Chapter 5 discusses the subjective testing that was performed; the purpose was to determine an appropriate number of equivalent sources to use to recreate a FEM-calculated sound field. Finally, Chapter 6 summarizes the results of the experimentation described in this thesis and makes a conclusion. Suggestions for future work in the virtual acoustic prototyping field are introduced.

Chapter 2: Auralization Theory

This chapter will discuss the theory that was used to develop the experiments discussed later in this thesis. Section 2.1 covers the production of 3-D sound via the use of head-related transfer functions. Section 2.2 illustrates the equivalent source method (ESM) in detail. Section 2.3 describes the numerical error associated with the ESM; the ultimate goal of this thesis is to determine the level of numerical error that can be detected by a listener during auralization.

2.1 Head-Related Transfer Functions

Head-related transfer functions (HRTFs) are measurements of the direction-dependent acoustic filtering that a free-field sound experiences due to the head, torso, and pinnae (outer ears). HRTFs are defined as a function of three variables: azimuth angle, θ, elevation angle, ϕ, and source frequency, ω. The definition of the azimuth and elevation angles is shown in Figure 2.1. They are defined as the angle between the line connecting the source and the center of the head and the line corresponding to zero azimuth and elevation. For a person in the standing position, zero azimuth and elevation is straight ahead and parallel to the ground.

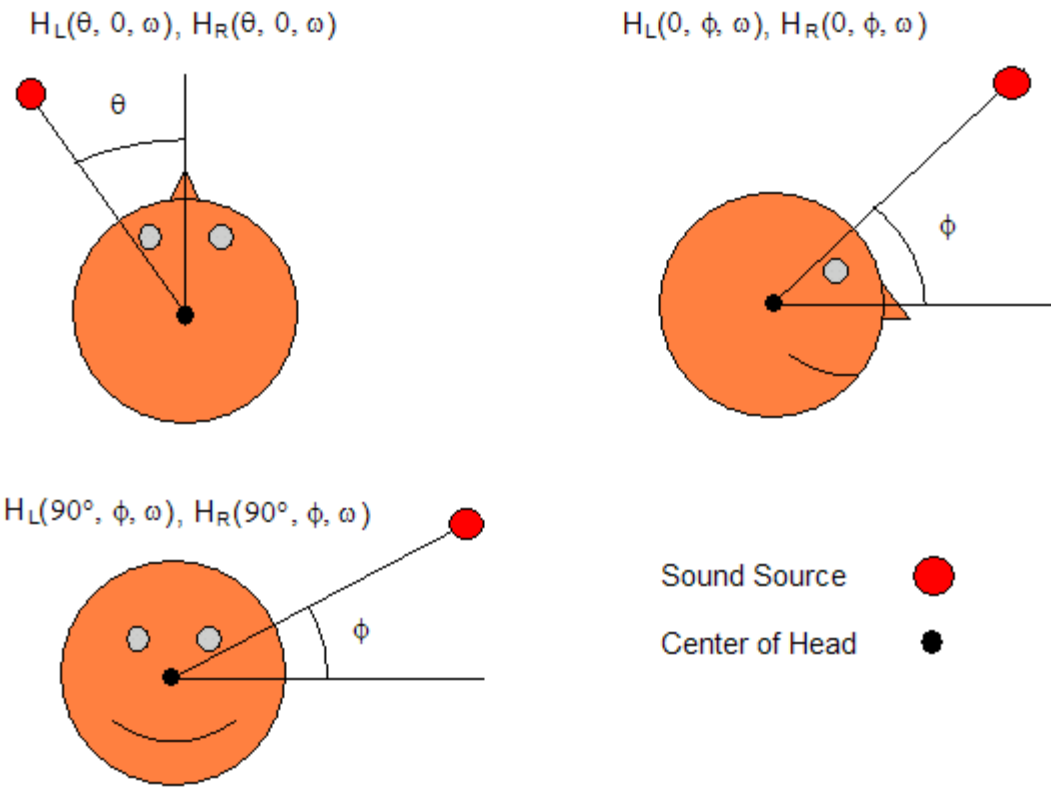

Figure 2.1. Definition of the HRTF azimuth angle, θ, and elevation angle, ϕ.

The simplest HRTF model – Duplex theory – was proposed by Rayleigh in 1907. [28]. Duplex theory states that our perception of sound direction depends on interaural time difference (ITD) and interaural intensity difference (IID). The head is modeled as a uniform sphere; impinging sound waves are diffracted and absorbed. The extent to which this occurs is highly dependent on the frequency of the sound source. The wave fronts arrive at each ear with different intensities and at different times. A low-frequency sound (with a wavelength much longer than the diameter of the head) originating from the left will first reach the left ear, and the wave will diffract around the head and arrive at the right ear microseconds later. A high-frequency sound (with a wavelength much smaller than the diameter of the head) is typically attenuated, or shadowed, by the head; since the wavelength is small. The IID is mainly a high-frequency effect, and the ITD is important at low frequencies.

Duplex theory has several limitations. It does not account for the shadowing of sound due to the pinnae and torso, and the absorption and diffraction of sound by the

head differs for each person. In addition, loci of sound source positions exist that can provide an identical ITD and IID [4]. Figure 2.2 shows one such locus of points, which is dubbed the "cone of confusion". According to Duplex theory, a sound source on the circle that describes the base of the cone (with an apex at the ear) is indistinguishable from a sound source at any other point on the circle. Although models such as Duplex theory is too simplistic to be effective, parameterization of HRTFs is still a current topic of research.

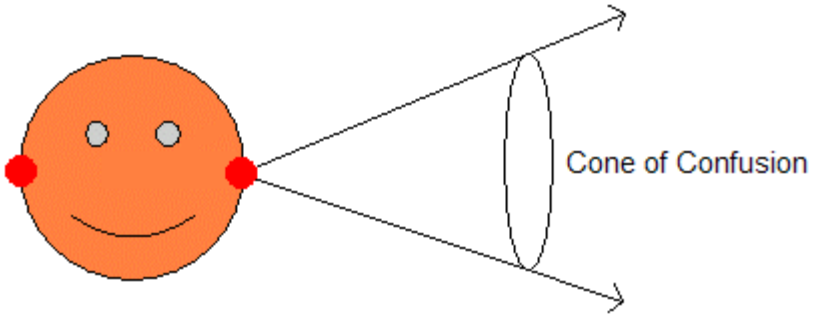

Figure 2.2. The "cone of confusion". According to Duplex theory, all points on the circle have the same ITD and IID.

A person's HRTFs are typically measured by inserting a tiny microphone into each ear canal. [4] A sound signal, which contains all frequencies of interest, is played over a loudspeaker at the elevation and azimuth angles of interest. The HRTF is the transfer function between the sound at the loudspeaker and the sound at the microphone. This cannot be measured directly; the transfer function between the loudspeaker input and microphone output is measured, and the dynamics of the loudspeaker and microphone are subsequently removed. In practice, HRTFs are easily measured for azimuth angles between -180° and 180°. HRTF measurement for elevation angles below -40° are usually not feasible since the ground impedes the placement of a sound source. However, sets of HRTFs that contain the full range of elevation and azimuth angles do exist.

2.2 The Equivalent Source Method and Least-Squares Solution

In a complex vibrating structure, every point on each moving surface acts as a sound source. Figure 2.3a shows a discretized vibrating plate (blue) that is in close proximity to a listener (black circle). Every blue point contributes to the total acoustic pressure at the black circle. Auralization of even a simple structure would be computationally infeasible because of the large number of sources. In an enclosure, reflections would also contribute to the pressure at the listener's location; reflections would need to be modeled as additional "virtual" sources. Intricate geometry and realistic surface types (absorptive or transmitting) would further complicate the determination of the sources. Due to their extreme number, it would be very difficult to apply HRTFs to all of these sources in order to conduct an auralization. However, finite element modeling (FEM) or the boundary element method (BEM) can calculate the acoustic pressure field in and around a structure.

ESM approximates a complex sound field, such as one resulting from the vibrating plate in Figure 2.3a, using a set of monopole sources placed on a sphere about the listener location, as shown in Figure 2.3b. The sound field in a small sphere – with a radius between 0.1 and 0.5 meters – can be accurately and efficiently recreated using a relatively small number of monopoles (shown in red).

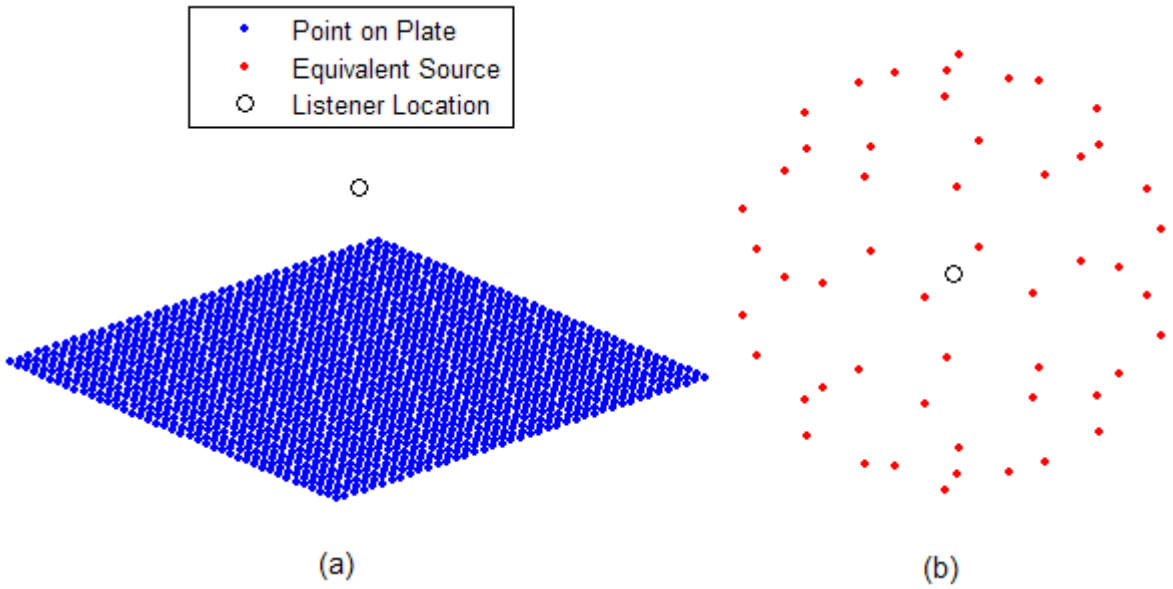

Figure 2.3. ESM approximates a complex sound field due to a vibrating structure (a) using a set of monopole sound sources placed on a sphere (b).

A monopole source is a radially-vibrating sphere that radiates sound equally in all directions. Figure 2.4 illustrates a monopole; the sphere produces sound by expanding and contracting along its radius. True monopoles only exist in theory, although a low-frequency loudspeaker attached to a speaker cabinet can closely approximate a monopole. [29] The sound propagation due to a monopole in the frequency domain is

$$p(R,\omega) = j\omega\left(\frac{\rho_0}{4\pi R}\right)Qe^{j\omega t}e^{-jkR} \tag{2.1}$$

where p is the acoustic pressure, R is the distance from the source to the receiver, ω is the sound frequency, ρ_0 is the density of air, and j equals $\sqrt{-1}$; k is the wave number and is equal to

$$k = \frac{\omega}{c} \tag{2.2}$$

where c is the speed of sound. The volume velocity Q is

$$Q = Su \tag{2.3}$$

where u is the velocity of the vibrating surface of the sphere, and S is the surface area of the sphere that is given by:

$$S = 4\pi \cdot r^2 \tag{2.4}$$

where r is the radius of the sphere.

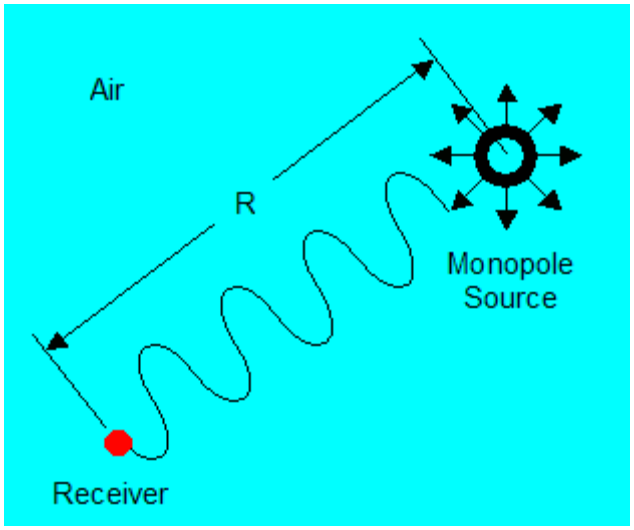

Figure 2.4. Propagation of sound due to a monopole source. The monopole radiates sound equally in all directions.

To carry out ESM, a structure is modeled using a finite element software package. The acoustic pressure response on a three-dimensional grid, or evaluation volume, around the listener location is calculated. The goal is then to determine the equivalent source strengths that recreate the FEM pressure field as closely as possible. The problem can be written in matrix form as:

$$T(\omega)\mathbf{Q}(\omega) \approx \mathbf{P}(\omega) \longrightarrow \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1m} \\ T_{21} & T_{22} & & T_{2m} \\ T_{31} & T_{32} & & T_{3m} \\ \vdots & & & \vdots \\ T_{n1} & T_{n2} & \cdots & T_{nm} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_m \end{bmatrix} \approx \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ P_n \end{bmatrix} \tag{2.5}$$

Here **P** is the FEM pressure field at *n* evaluation points, and **Q** is the vector of volume velocities for *m* equivalent sources. The matrix *T* contains the $m \cdot n$ transfer functions that represent the contribution of the sources to the acoustic pressure at the evaluation points. For example, the transfer function T_{32} is:

$$T_{32}(\omega) = j\omega\left(\frac{\rho_0}{4\pi R_{32}}\right)e^{-jkR_{32}} \tag{2.6}$$

where R_{32} is the distance between the second source and the third evaluation point. T_{32} represents the contribution of the second source to the acoustic pressure at the third node. Note that

$$T_{32}(\omega) \neq \frac{P_3(\omega)}{Q_2(\omega)} \tag{2.7}$$

but rather

$$P_3(\omega) = T_{31}(\omega)Q_1(\omega) + T_{32}(\omega)Q(\omega)_2 + \ldots + T_{3m}(\omega)Q_m(\omega) \tag{2.8}$$

Linearity and the principle of superposition is assumed when combining the equivalent sources to recreate the approximate FEM pressure field. The least-squares method (LSQR) [30] is used to provide an approximate solution to equations of the form

$$A\mathbf{x} = \mathbf{B} \tag{2.9}$$

LSQR, an iterative method which continually seeks a solution with lower error, is not the only way to solve the system; it was chosen because it provides good conditioning properties. In the case of ESM, the goal is to minimize the LSQR error, e_{LSQR}, which is equal to the root-mean-square difference in the acoustic pressure at all *z* nodes in the evaluation volume. The error is between the model and recreated sound field:

$$e_{LSQR} = \sqrt{\sum_{i=1}^{z}\left(T_{i1}(\omega)Q_1(\omega) + T_{i2}(\omega)Q_2(\omega) + \cdots + T_{im}(\omega)Q_m(\omega) - P_i(\omega)\right)^2} \tag{2.10}$$

To obtain the equivalent source strengths, Equation 2.10 is solved once for each frequency in the band of interest.

2.3 Numerical Error

The virtual acoustic prototyping process described in Section 1.1 cannot exactly reproduce the binaural signal that will be present at a listener's ears in the physical structure. There are four major sources of error associated with an auralization:

- Discrepancies between the finite element model and physical structure
- Numerical error is the solution of the finite element model
- Reproduction of the calculated sound field in a small evaluation volume
- Use of a limited number of equivalent sources in LSQR algorithm
- Interpolation and approximation in the application of HRTFs

The error from the early steps propagates through to subsequent steps. For this reason, it is imperative that the finite element model accurately describes the structure. In this thesis, the last three sources of error in the above bulleted list are of primary interest; they will be quantified and subjectively evaluated.

The modeling of the structure's acoustics is the stage in the auralization process that is most sensitive to input parameters. The calculated acoustic pressure field that serves as input to the ESM software depends on a variety of specifications in the model. In order to have an accurate output, the properties of the materials that are used in software must match those of the structure's building materials. In addition, the assembly constraints and boundary conditions that are used in software must match the assembly conditions of the final structure. Materials such as plywood and medium-density fiberboard (MDF) are difficult to characterize because no two batches are exactly alike. When two wooden panels are bolted together, internal stresses are introduced into the material; this phenomenon is nearly impossible to account for using a commercial finite element package. The natural frequencies of a vibrating structure can usually be determined with an error of 2-5%. In acoustic modeling, FEM can typically provide an

accurate solution at low frequencies (less than 500 Hz), but at higher frequencies, obtaining an accurate solution becomes computationally infeasible. The frequency range of interest in this thesis is between 50 and 1500 Hz. The analysis in Chapter 4 will show that this step is the source of error in the auralization performed in this thesis. For this reason, the finite element model accurately describe the structure. Matching the finite element model output to the acoustic response of a physical prototype is an extremely difficult task. However, it is possible to use FEM to create a model that is representative of the prototype. This means that although the numerical correlation between the model output and prototype response will be low, the sounds of the model output and prototype response will be very similar and will have the same psychological effect on the listener. Section 4.2 will discuss this issue further; in this thesis high numerical correlation was not achieved between the modeled structure and the physical structure, but the sounds of their responses were remarkably similar.

ESM recreates the sound field that is calculated by the finite element software; instead of using the entire model output, only the output in a small evaluation volume is used. The evaluation volume is a sphere that is centered around the probable location of the listener's head. A set of nodes from the finite element model's acoustic space must be contained within the evaluation sphere. The accuracy of an auralization depends on the radius of the sphere. Theoretically, the sound "field" inside of an evaluation sphere containing only one node can be exactly reproduced using only one equivalent source. However, the binaural signal will not sound like a vibrating structure; it will sound like a single concentrated source located somewhere outside of the head. To create an accurate sound field using equivalent sources, the evaluation sphere should be as large as possible. The available computational power will limit the number of equivalent sources – thus limiting the size of the sphere that can be auralized.

The accuracy of ESM depends on the number of evaluation points that are used in the calculation of the equivalent source strengths. The number of evaluation points *NEP* is equal to

$$NEP = \frac{S_e}{\left(\lambda_{\min} / EPWL\right)^2} \tag{2.11}$$

where *EPWL* is the "evaluation points per acoustic wavelength" parameter, λ_{min} is the smallest acoustic wavelength (which corresponds to the highest frequency of interest), and S_e is the surface area of the evaluation sphere given by

$$S_e = 4\pi \cdot r_e^2 \tag{2.12}$$

where r_e is the evaluation sphere radius. Not all of the nodes in the evaluation volume are used to calculate the equivalent source strengths; Figure 2.5 illustrates the relationship between finite element model nodes, evaluation points, and equivalent sources. From the locus of nodes in the evaluation sphere (green and red), a certain number are chosen to be evaluation points (red). Equation 2.11 specifies the number of nodes that should be used as evaluation points. The point of diminishing return for using more evaluation points is near *EPWL* = 6. [31]

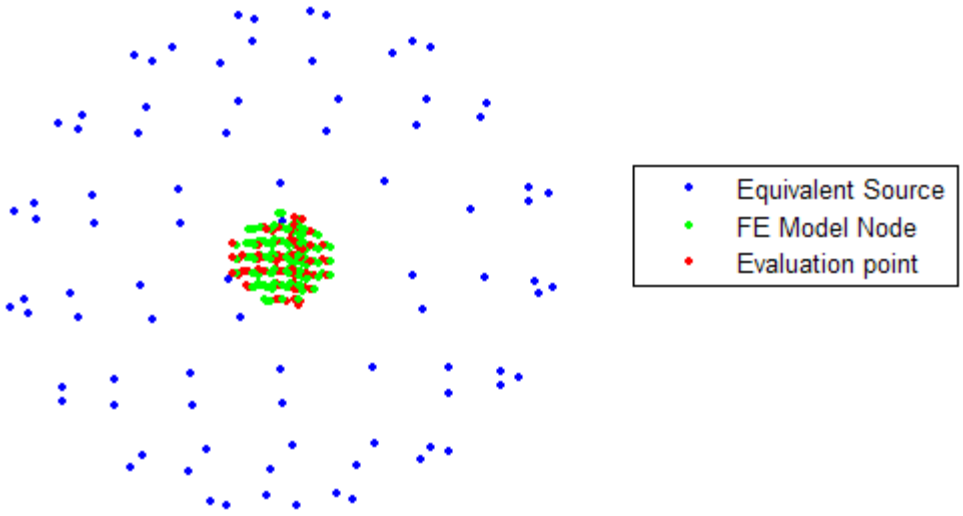

Figure 2.5. Finite element (FE) model nodes, evaluation points, and equivalent sources. A certain number of nodes are chosen as evaluation points, which are used to calculate the equivalent sources.

The number of equivalent sources that are used also depends on the size of the evaluation sphere. The number of sources *NES* is given by

$$NES = \frac{S_e}{\left(\lambda_{\min} / SPWL\right)^2} \tag{2.13}$$

where *SPWL* is the "sources per acoustic wavelength" parameter. In theory, the point of diminishing return for using more evaluation points is near *SPWL* = 3.7. [31] All sources are placed on a sphere that is concentric with the evaluation sphere; the source sphere has a minimum radius of 0.7 meters greater than the radius of the evaluation sphere. Equation 2.6 is used to calculated the matrix of transfer functions *T* between the evaluation point acoustic pressures **P** and the source strengths **Q**. The matrix *T* should not be square; if the number of sources is equal to the number of evaluation points, then the pressures at the evaluation points can be reproduced without error, but the error at the other nodes in the evaluation sphere would be high (in mathematical terms, this is known as ill-conditioning). In addition, the calculated source strengths could be excessively large. Since the acoustic pressure values at the nodes in the evaluation sphere is recreated using a smaller number of equivalent sources, there is no exact solution to

$$T\mathbf{Q} = \mathbf{P} \tag{2.14}$$

so the least-squares (LSQR) algorithm is used to solve Equation 2.5 and reduce the error (which is defined in Equation 2.10) to an acceptable level. Other algorithms were considered, but LSQR was found to be the most efficient. [31] Figure 2.6 shows the error associated with the LSQR algorithm as a function of evaluation sphere radius and number of equivalent sources. [32] A diffuse field containing four different frequencies was recreated using ESM. the lowest error occurs when the evaluation sphere is small and the number of equivalent sources is large. However, the use of a small evaluation sphere will not necessarily produce an accurate result. The LSQR algorithm has a greater error at higher frequencies because the acoustic wavelengths are shorter (and thus *SPWL* is lower).

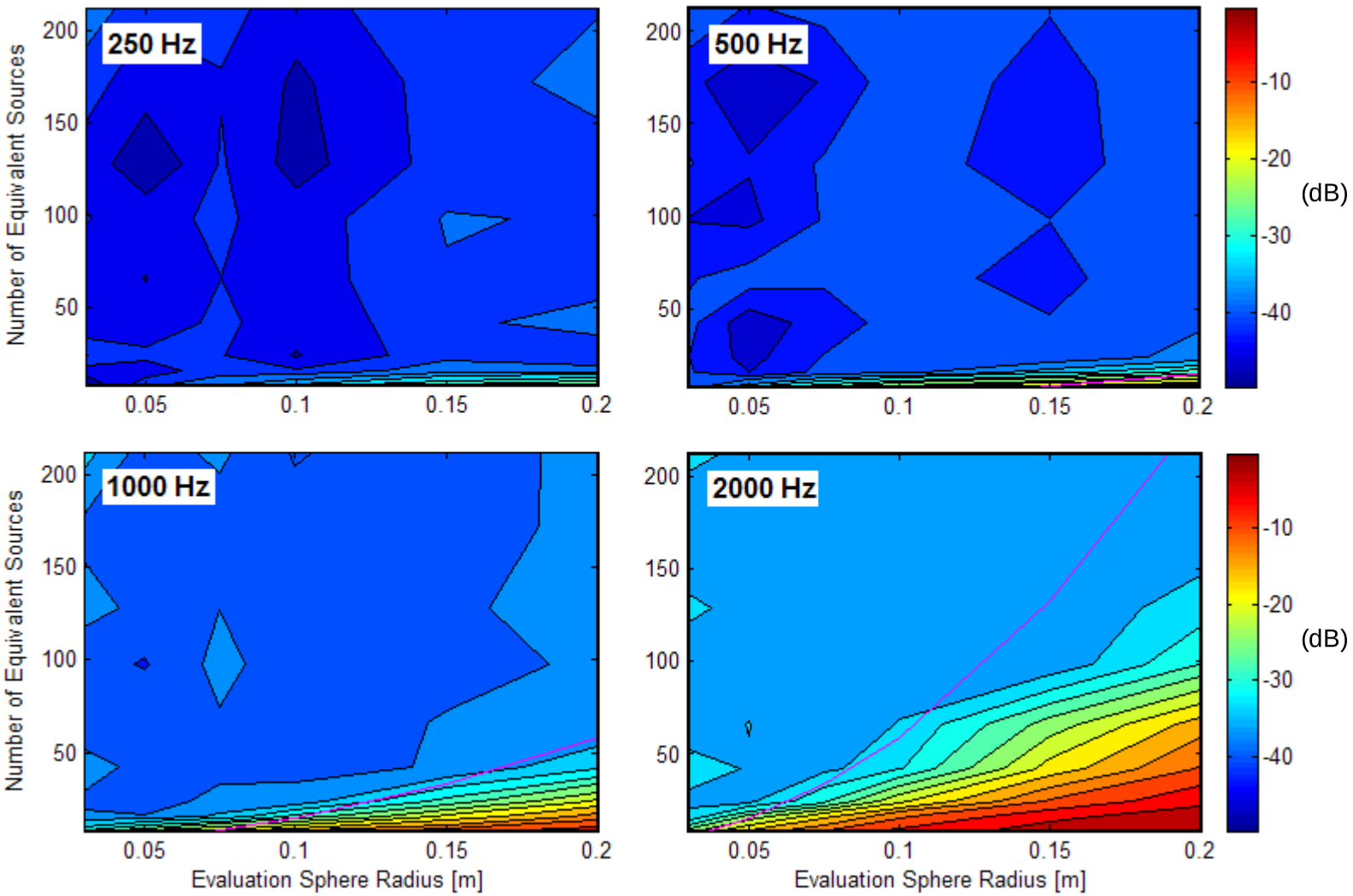

Figure 2.6. LSQR error as a function of evaluation sphere radius and number of equivalent sources. The magenta line indicates 3.7 sources per wavelength.

The application of HRTFs to the equivalent sources produces the binaural signal that is played over a set of headphones; this step introduces additional error to the auralization process. Each person's head is different, which means no two people have exactly the same HRTFs. Often a representative set of HRTFs is used, such as those of the Knowles Electronic Manikin for Acoustics Research (KEMAR). [33] Another source of error comes from the interpolation and extrapolation of HRTFs. HRTFs are usually measured for all azimuth angles at 5° intervals. Due to physical constraints, they are usually only measured for elevation angles between -40° and 90° at 5° intervals. ESM places sources on the source sphere at all elevation and azimuth angles. If a source is placed at an azimuth of 56° and an elevation of 30°, then the appropriate HRTF must be interpolated from the known HRTFs in the set. The 30-degree-elevation HRTFs for azimuth angles of 55° and 60° are combined in a weighted average to provide an approximation of the 56° HRTF. For sources that lie below -40°, HRTFs are extrapolated. [34] At locations closer to the north and south poles (elevation = 90° and –90°, respectively), the HRTFs' magnitude versus frequency curves tend to flatten because pinnae interactions are less prevalent. The south pole HRTF is calculated by averaging the -40-degree-elevation HRTFs for all azimuth angles. The HRTFs for elevations between -40° and -90° are then calculated by interpolation. Figure 2.7 shows the average error in the HRTF application process as a function of number of equivalent sources and frequency. [32] At frequencies less than 1500 Hz, HRTF approximation error is miniscule. Above 3000 Hz, the HRTF error is more or less constant across the frequency band. The error tends to diminish as more equivalent sources are included in the auralization.

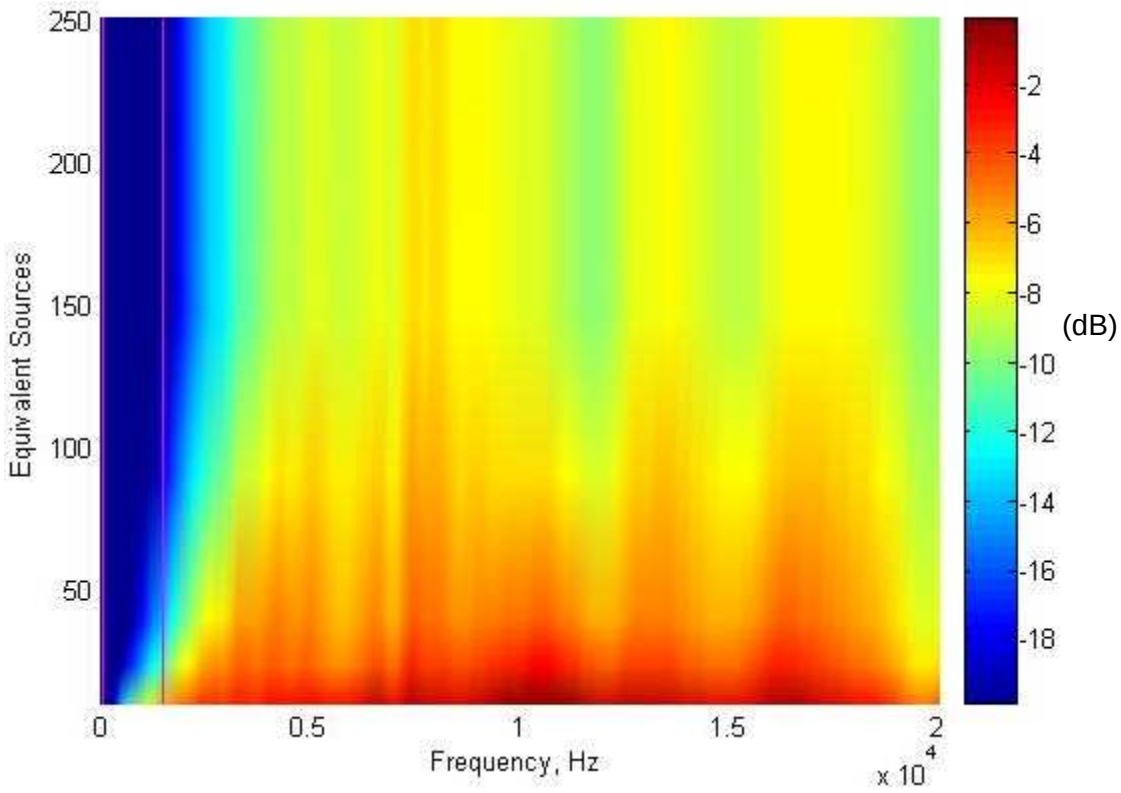

Figure 2.7. HRTF application in the frequency range of human hearing. The magenta lines indicate the band from 50 to 1500 Hz, which is of interest in this thesis.

Figure 2.8 shows the combined ESM and HRTF application error as a function of evaluation sphere radius and number of equivalent sources. [32] This is the total error associated with the virtual acoustic prototyping process, not including inaccuracies in the finite element model. It is the difference between the binaural signal created from a set equivalent sources and the binaural signal created from thousands of original sources (such as the blue dots shown in Figure 2.3). The total error is lowest when a relatively large evaluation sphere and many equivalent sources are used. Error increases with frequency because the LSQR algorithm error is greater. Figure 2.8 indicates that a limit on the number of equivalent sources is present in an auralization system, then an upper limit on the useful frequency range of ESM may exist. At 2000 Hz, the numerical error exceeds -15 dB for the entire feasible range of evaluation sphere sizes and equivalent source quantities. Chapter 5 will subjectively determine what numerical error level is noticeable in several different kinds of sounds. The virtual acoustic prototyping method must be able to keep the numerical error below this perceptual threshold.

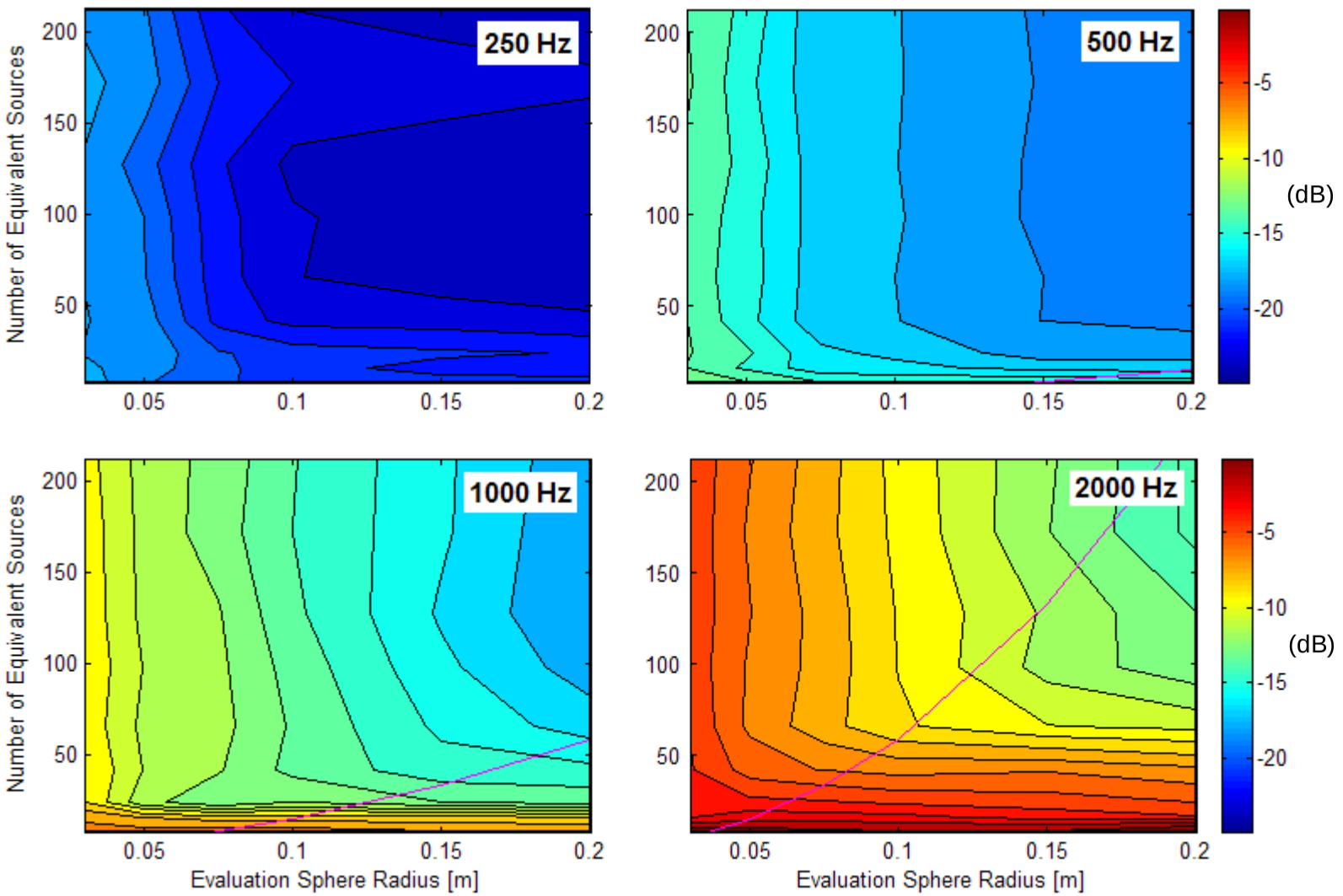

Figure 2.8. Total ESM error plus HRTF application error. The magenta line indicates 3.7 sources per wavelength.

Chapter 3: Numerical Modeling

This chapter will describe the numerical modeling of the virtual acoustic prototyping process for a hypothetical structure. This modeling performed here will provide the structural acoustic data set that is necessary to carry out the virtual acoustic prototyping process described in Section 1.1. Section 3.1 briefly describes the finite element method (FEM) and its applications. The application of FEM to a rod that is subject to an axial force is shown. Section 3.2 covers the finite element modeling that was necessary for calculating the structural acoustic response. Section 3.3 explains the application of the equivalent source method (ESM) and head-related transfer functions (HRTFs) to the finite element model output.

3.1 FEM Theory

The finite element method is an engineering analysis tool that can be applied to structures of complex geometry. An intricate structure is discretized into finite elements of simple geometry, and the laws of physics are applied to each element to form a coupled system of equations. The structural response is calculated by solving the system of equations. FEM has application in many engineering problems, including solid mechanics, fluid mechanics, heat transfer, and structural acoustics. [35]

Figure 3.1 shows three example applications of FEM. Figure 3.1a is a model an I-beam. The structural response – which can include natural frequencies, mode shapes, stress, and displacement – is solved for a specified set of material properties, boundary conditions, and force inputs. Figure 3.1b is a model of the fluid flow around the cross-section of an automobile body. The lift and drag forces on the car are calculated for a specified set of material properties and flow characteristics. Figure 3.1c is a model of the complex geometry of a disc-brake rotor. The temperature distribution is calculated for a given set of air and rotor properties, initial conditions, and heat inputs.

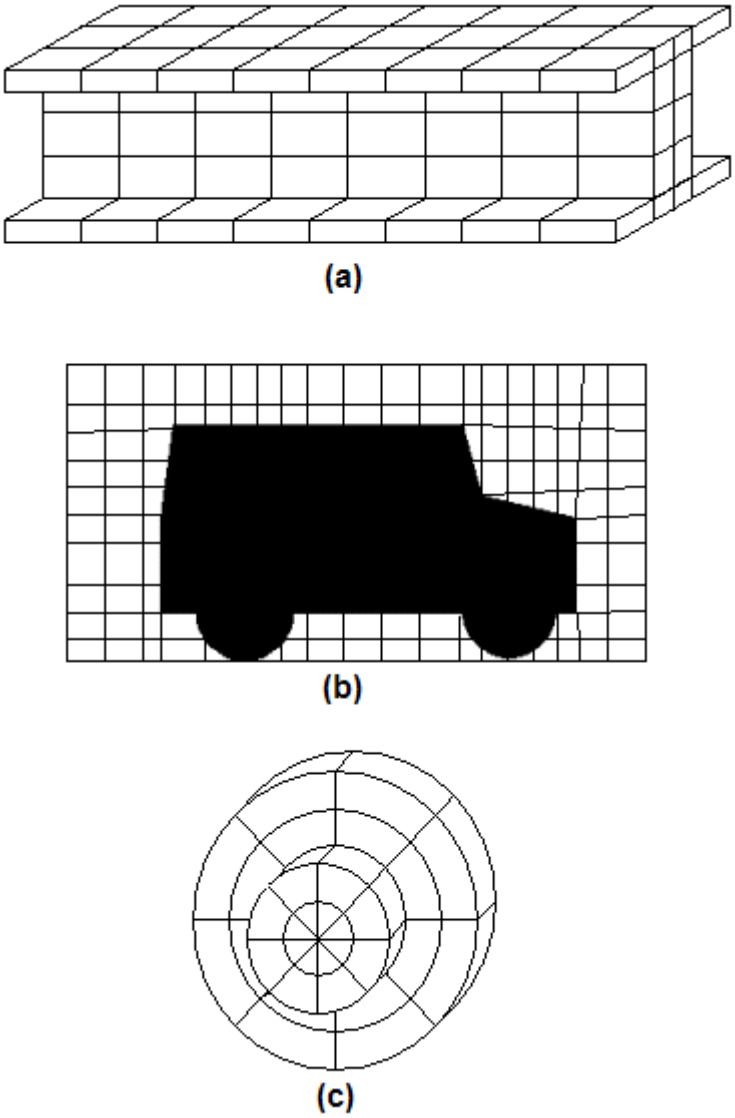

Figure 3.1. Several applications of FEM: (a) mechanical model of an I-beam (b) air flow model over an automobile body (c) heat transfer model of a brake rotor.

FEM can also be used to model the interaction between a vibrating solid and an acoustic medium. Figure 3.2 shows a two-dimensional model of a thin plate acting on an acoustic space with rigid walls. A complex geometry was chosen because no analytical exists, and a numerical method such as FEM must be used to solve the model. The plate is excited at a particular node or node set by a mechanical force of the plate. The plate and acoustic space are constrained in a way such that as the plate deforms, the air behind the plate moves as well. The vibration of the plate causes the propagation of a pressure wave, or sound. The reflection of the sound off of the rigid surfaces results in a complex

pressure field inside of the acoustic space. Using finite element analysis software such as ABAQUS, the transfer function (for a discrete set of frequencies) between the input force and acoustic pressure can be calculated for each node.

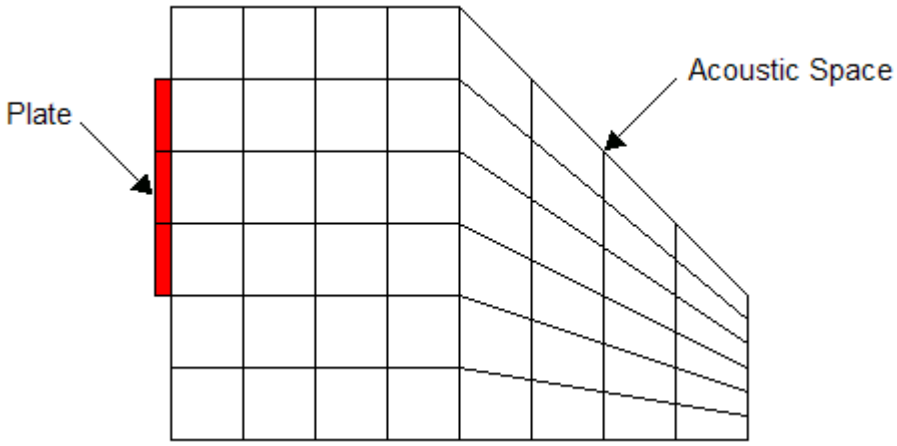

Figure 3.2. Two-dimensional finite element model of a thin plate acting on an acoustic space.

The simplest application of FEM is a rod that is subject to an axial force. Figure 3.3a shows a rod that is fixed at one end; the other end experiences a compressive force *P*. For this analysis, the rod will have mass *m*, length *L*, cross-sectional area *A*, and modulus of elasticity *E*. The bending of the rod under its own weight is assumed to be negligible. In Figure 3.3b, the rod is discretized into six masses m_i that are connected by five springs k_i. The finite element representation of the rod is shown in Figure 3.3c; the model consists of five elements that connect six nodes. The equation of motion for a multiple-degree-of-freedom (MDOF) spring-mass-damper system is:

$$M\ddot{\mathbf{x}} + C\dot{\mathbf{x}} + K\mathbf{x} = \mathbf{F} \tag{3.1}$$

where *M*, *C*, and *K* are the mass, damping, and stiffness matrices, respectively, **F** is the force vector, **x** is the displacement vector, and $\ddot{\mathbf{x}}$ is the acceleration vector.

Expanding Equation 3.1 for the system of Figure 3.3b (with $C = 0$) results in:

$$\begin{bmatrix} m_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & m_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & m_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & m_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & m_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & m_6 \end{bmatrix} \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \\ \ddot{x}_4 \\ \ddot{x}_5 \\ \ddot{x}_6 \end{bmatrix} +$$

$$\begin{bmatrix} k_1 & -k_1 & 0 & 0 & 0 & 0 \\ -k_1 & k_1+k_2 & -k_2 & 0 & 0 & 0 \\ 0 & -k_2 & k_2+k_3 & -k_3 & 0 & 0 \\ 0 & 0 & -k_3 & k_3+k_4 & -k_4 & 0 \\ 0 & 0 & 0 & -k_4 & k_4+k_5 & -k_5 \\ 0 & 0 & 0 & 0 & -k_5 & k_5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -P \end{bmatrix} \qquad (3.2)$$

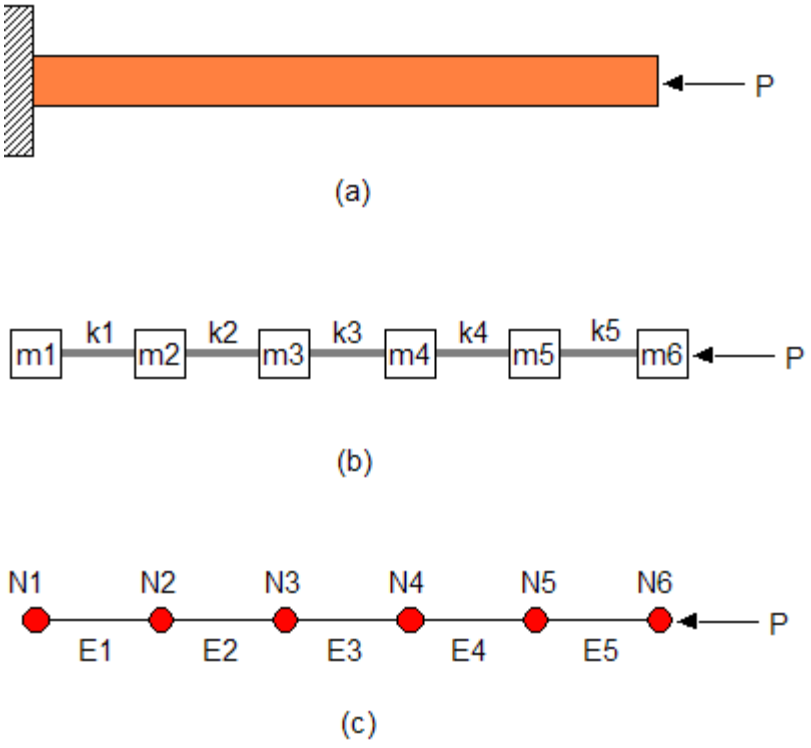

Figure 3.3. Application of FEM to a rod in axial compression. (a) Rod fixed at one end and forced at the other end. (b) Discretization of the rod into a six-degree-of-freedom system. (c) Finite element representation of the rod using six nodes and five elements.

For a rod in axial compression, the longitudinal deflection δ relates to the dimensions, modulus of elasticity, and force magnitude according to:

$$\delta = \frac{PL}{AE} \tag{3.3}$$

Therefore, the stiffness k is:

$$k = \frac{P}{\delta} = \frac{P}{\left(\frac{PL}{AE}\right)} = \frac{AE}{L} \tag{3.4}$$

Since the rod is uniform along its length, each stiffness k_i is equal to:

$$k_i = \frac{A_i E}{L_i} = \frac{AE}{(L/4)} = \frac{4AE}{L} = k \tag{3.5}$$

The masses m_i are determined for each node by taking the mass between node i and node $i+1$ and dividing it evenly between node i and node $i+1$. Masses m_1 and m_6 are equal to $0.1m$, and masses m_2 through m_5 are equal to $0.2m$. The resulting equation of motion for the system is given by:

$$\begin{bmatrix} 0.1m & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.2m & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2m & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.2m & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.2m & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.1m \end{bmatrix} \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \\ \ddot{x}_4 \\ \ddot{x}_5 \\ \ddot{x}_6 \end{bmatrix} +$$

$$\begin{bmatrix} k & -k & 0 & 0 & 0 & 0 \\ -k & 2k & -k & 0 & 0 & 0 \\ 0 & -k & 2k & -k & 0 & 0 \\ 0 & 0 & -k & 2k & -k & 0 \\ 0 & 0 & 0 & -k & 2k & -k \\ 0 & 0 & 0 & 0 & -k & k \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -P \end{bmatrix} \tag{3.6}$$

The fixed end condition is modeled by letting $x_1 = 0$; the first equation is eliminated and the system becomes five-degree-of-freedom. This example is a simple case that can be solved without special software. Finite element software packages such as ABAQUS, NASTRAN, and ANSYS are useful tools for formulating and solving models with complex geometry, material properties, and loading conditions.

3.2 Calculating the Structural Acoustic Response

Figure 3.4 shows the design of a structural acoustic test platform that was used as a tool to validate the ESM software. The enclosed interior space is a rectangular box (44 ½" x 49" x 69 ¼") minus a triangular section as shown. The walls are single-ply, ¾" medium-density fiberboard (MDF), and the panel that contains the window is double-ply. The window is ¼" Lexan with a 22" x 36" visible area; all edges are clamped securely. The enclosure was designed to be constructed in six sections. The panels overhang each other so all bolted connections can be made from the outside. Since the inside walls of the enclosure are smooth, the finite element modeling of the interior space was simpler. A flush-mounted door in the side allows for access to the inside. Excitation of the acoustic space can occur either by driving the an interior speaker or the window.

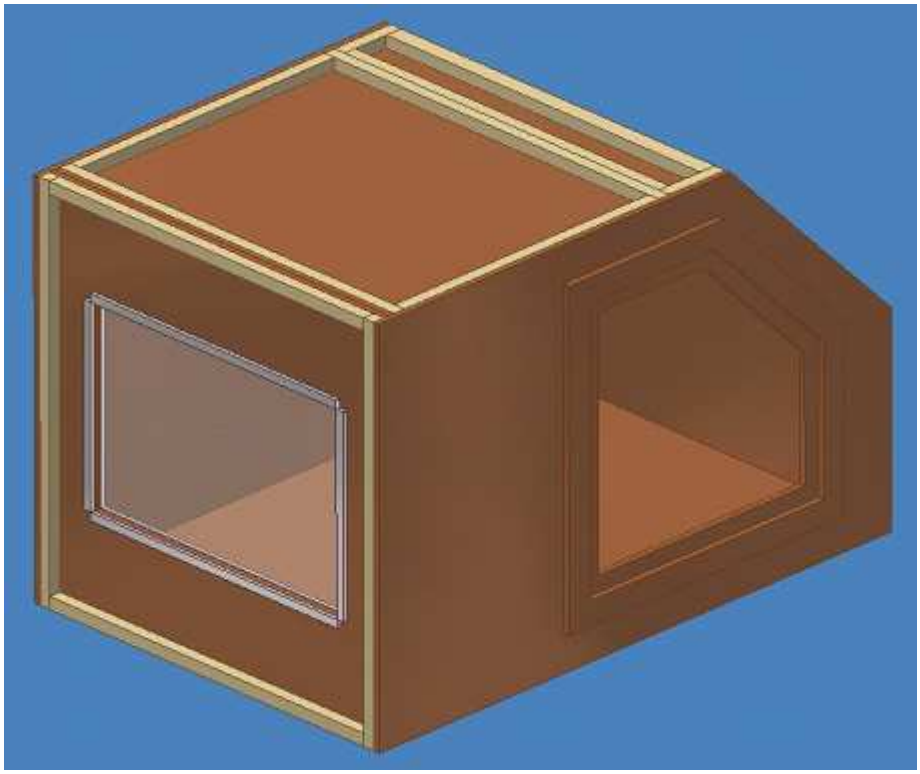

Figure 3.4. CAD model of a structural acoustic test platform.

The theoretical response of the structure acoustic test platform was calculated using ABAQUS – a FEM software package. The accuracy of the model output relies on the model parameters, so great care was taken when choosing how to define the enclosure in software.

Establishing the geometry of the panels was the first step in the FEM process. For simplicity, the acoustic space and each panel was modeled, but the trim pieces and bolted joints were neglected and treated as boundary conditions in a later stage. The MDF panels were sized such that the perimeter of each panel was equal to the perimeter of the corresponding face of the acoustic space. Figure 3.5 shows the enclosure assembly; the panel faces were mated to the faces of the acoustic space (which is obscured by the panels in the picture). The windowed panel is double-ply, so the inner MDF panel was constrained to both the acoustic space and the outer MDF panel. The window hole in the inner panel was designed to be two inches longer in width and height, such that the Lexan window could be inset and bolted to the outer MDF panel. In the model, the perimeter of

the Lexan plate was fixed to the outer panel. No constraint was established between the window and the inner panel.

Figure 3.5. Assembly of the enclosure. The Lexan and MDF panels were sized around the acoustic space (hidden by panels), and the panel faces were mated to the faces of the acoustic space.

The material properties that are used in the finite element model have a direct effect on the natural frequencies and damping ratios of the acoustic and structural modes of the enclosure. The widely-published value for the bulk modulus of air E is 140,000 N/m^2, and the density of air is obtained handily from the ideal gas law:

$$PV = mRT \quad \longrightarrow \quad \rho = \frac{m}{V} = \frac{P}{RT} \tag{3.7}$$

where P is the air pressure, T is the air temperature, and R = 287 J/kg-K is the gas constant for air. The density of air was set to 1.15 kg/m^3 in ABAQUS; this is slightly

lower than many published values because of the altitude – and thus lower air pressure – of Blacksburg, Virginia. Lexan is generally manufactured to a tight tolerance; MDF varies greatly between brands and production runs, but measurement of material properties for each panel would have been impractical. In light of this, average published values of modulus of elasticity E and Poisson's ratio v were used for Lexan and MDF. Table 3.1 summarizes the material properties that were used in the enclosure model. To calculate the internal damping of the Lexan and MDF, proportional damping was assumed:

$$C = \alpha M + \beta K \tag{3.8}$$

where α and β are the proportional damping constants. The rational fraction polynomial method of modal analysis [36] was performed to determine the first ten dominant natural frequencies ω_i and their corresponding damping ratios ζ_i. Using a pair of natural frequencies and damping ratios, the damping constants can be calculated according to

$$\zeta_i = \frac{\alpha}{2\omega_i} + \frac{\beta\omega_i}{2} \qquad i = 1,2,3,\ldots,n \quad [37] \tag{3.9}$$

Ten different pairings of i values were used to obtain an average estimate of α and β.

Table 3.1. Material properties for Lexan, MDF, and air.

Property	Lexan	MDF	Air	Units
α	0.000006	0.00012	N/A	none
β	0.5	0.5	N/A	none
ρ	1075	800	1.15	kg/m^3
E	900	3500	0.14	MPa
v	0.37	0.37	N/A	none

Part interactions, boundary conditions, and forcing functions were modeled after the enclosure geometry was established. The panels were allowed to breathe as the acoustic pressure inside changed. The displacements of the air and panels at the edge of the acoustic space were tied together. As a panel vibrates, the air touching it vibrates as

well; this phenomenon causes the propagation of sound waves. The translational displacement of the enclosure was constrained in space via the use of boundary conditions. All edges of the acoustic space – where the MDF panels join together – were fixed in the x-, y-, and z-directions. The edges were also not allowed to rotate; this simulates the bolted connections on the structural acoustic test platform that was built. Excitation of the acoustic space was achieved by a sinusoidal force (at the frequency of interest) on the window at a location four inches left and down from the top-right corner of the window. Figure 3.6 illustrates the boundary conditions and loads for the enclosure model.

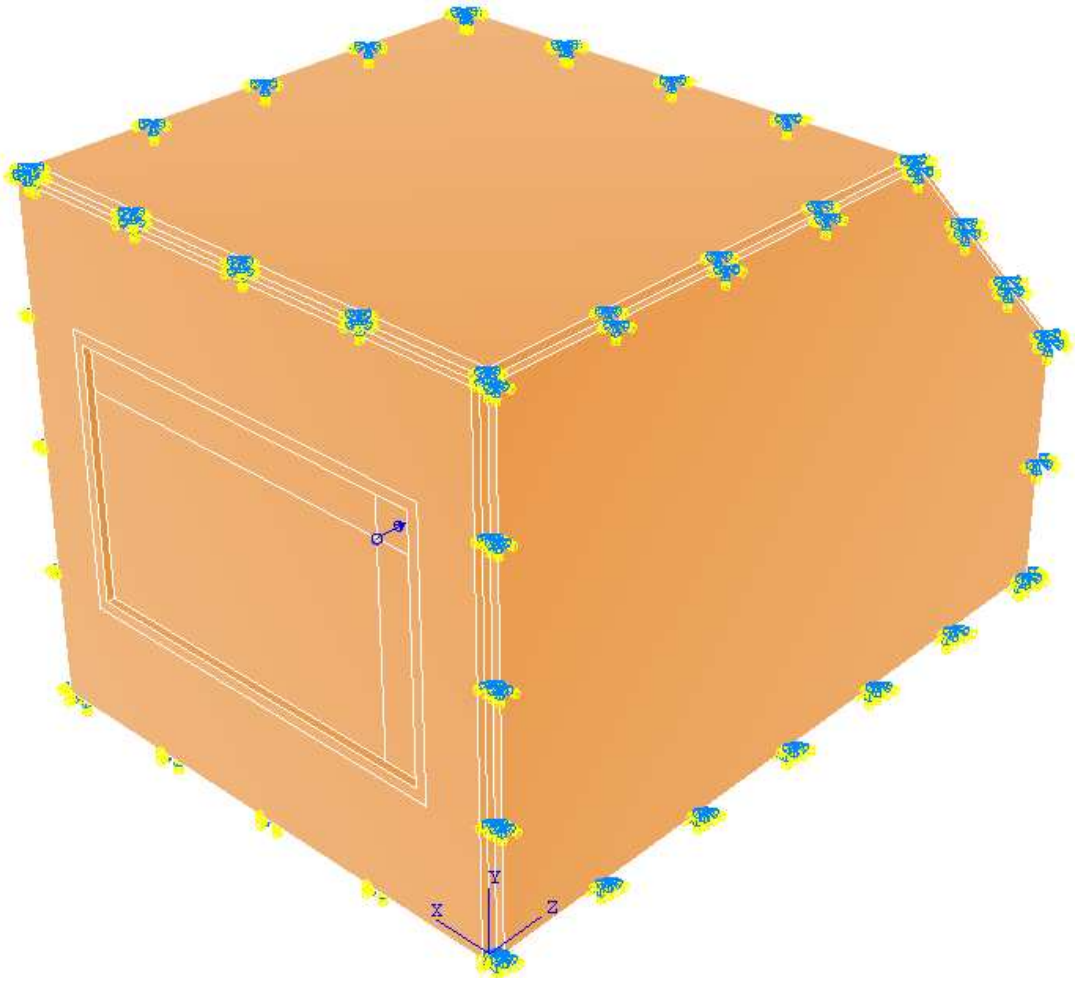

Figure 3.6. Loads and boundary conditions for the enclosure model. The blue arrow on the window shows the location of the sinusoidal force input. The blue and yellow arrows indicate the edges that were constrained in all six (three translational and three rotational) degrees of freedom.

The final step, before solving, in the FEM process is meshing. Figure 3.7 shows the mesh for the structural acoustic test platform. The general rule of thumb for finite element models in acoustics is to used at least six elements per wavelength. Wavelength λ is related to frequency f by the speed of sound c:

$$\lambda = \frac{c}{f} \tag{3.10}$$

Element with a characteristic length of two inches satisfy this rule of thumb for frequencies up to 1145 Hz.

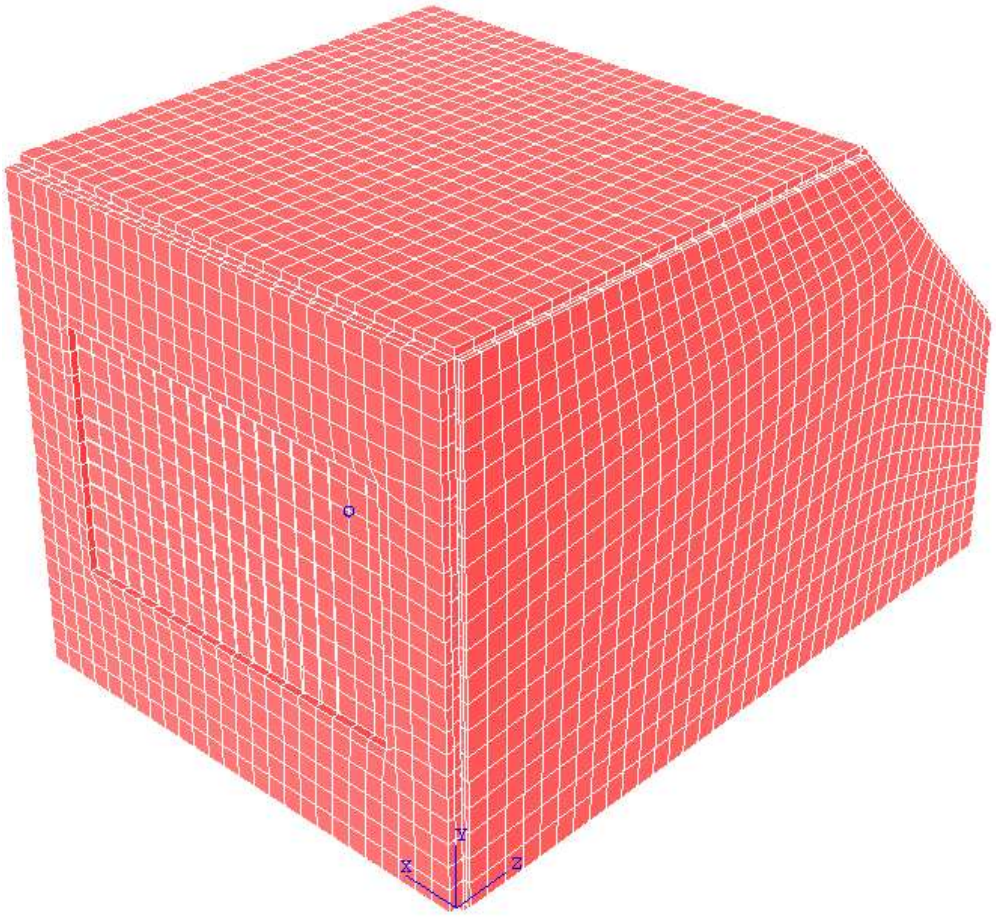

Figure 3.7. Mesh of the enclosure model. The characteristic length of a typical element was two inches, which provides more than six elements per wavelength.

The response of the acoustic space was calculated at every integer frequency between 50 and 1500 Hz. Above 1145 Hz, the accuracy may be slightly reduced because fewer than

six elements per wavelength were used. Since the impulse response is of interest, the sinusoidal response of the acoustic space was later divided by $\omega/((j\omega)^2+\omega^2)$ at each frequency; the result is the response of the acoustic space to an impulse force on the window. ESM and HRTFs are applied to the finite element model output in Section 3.3. The model response is compared to the constructed structural acoustic test platform in Chapter 4.

3.3 Application of ESM and HRTFs

After FEM, the next step in the auralization process is the application of ESM to an evaluation volume within the modeled acoustic space. Figure 3.8 shows the location of the head and evaluation spheres that were specified. The head directly faces the forcing location on the window and is located slightly off-center is all three directions. The cyan circles indicate the radii of the evaluation spheres (r_e = 0.14, 0.17, 0.20, and 0.23 meters). The sphere sizes were picked because Figures 2.5 and 2.7 show evidence of a numerical error gradient within the chosen range. In addition, smaller spheres were not feasible because the finite element mesh was not fine enough. ESM calls for a certain number of evaluation points per acoustic wavelength (*EPWL*, refer to section 2.3), and small spheres do not contain a sufficient number of nodes to meet this requirement. A finer mesh would allow smaller evaluation spheres to be evaluated, but this was not feasible due to the size of the structure and the available computational power.

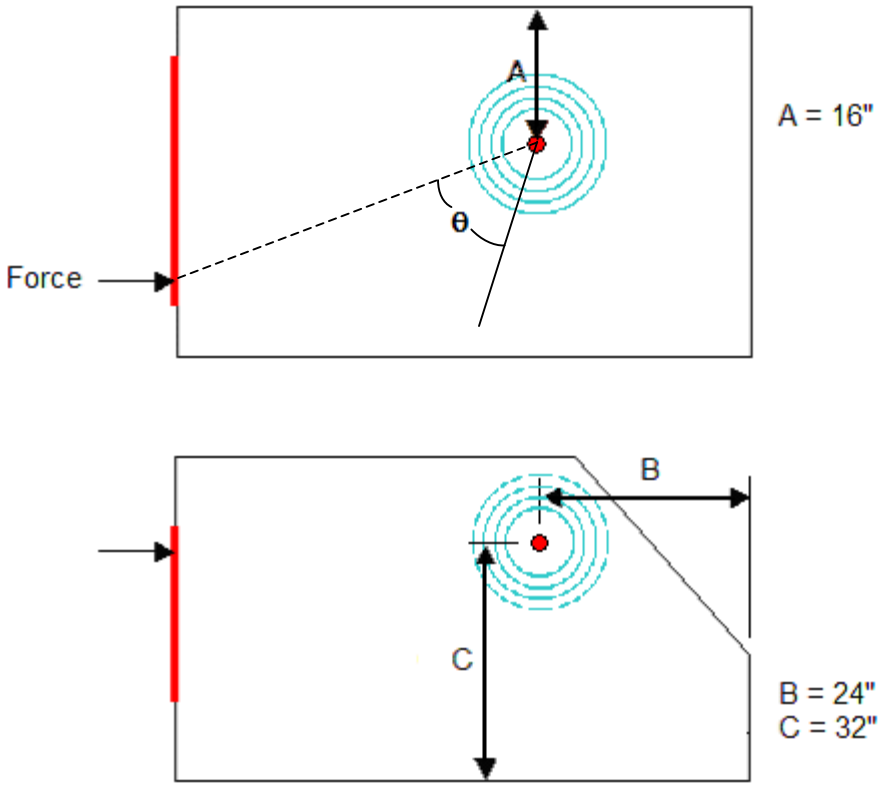

Figure 3.8. Location of the head and evaluation spheres. ESM was applied to the finite element model output using four sizes of evaluation sphere (shown in blue). The head location is shown by the red dot.

The eight equivalent source quantities, or quality levels, that were chosen for each sphere size correspond to 1.5, 2.0, 2.5, 3.0, 3.5, 3.7, 4.0, and 4.5 equivalent sources per acoustic wavelength (*SPWL*, refer to Section 2.3). Prior research indicates a diminishing return on using more equivalent sources at approximately *SPWL* = 3.7. [31] Figure 3.9 illustrates the eight equivalent source recreations of the acoustic pressure field in the four evaluation volumes (yielding a total of 32 cases). The highest quality level corresponds to *SPWL* = 4.5, or 258 equivalent sources; *SPWL* should not meet or exceed *EPWL* or ill-conditioning problems will occur.

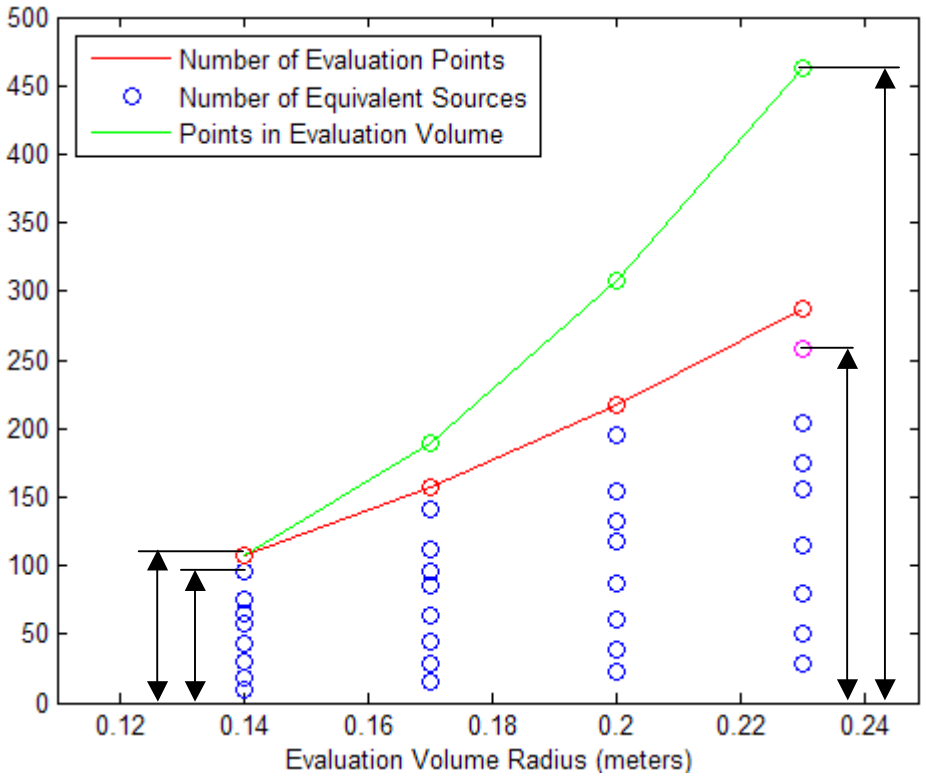

Figure 3.9. Evaluation sphere sizes and quality levels. ESM was performed at 32 different error levels (magenta and blue circles), with the magenta circle indicating the most accurate case. The number of equivalent sources and the number of evaluation points (red line) were not allowed to exceed the number of nodes in the evaluation volume (green line).

The numerical least-squares (LSQR) error associated with ESM is shown in Figures 3.10, 3.11, and 3.12. Figure 3.10 includes four plots – one for each of four different *SPWL* values – indicating that as the evaluation sphere radius increases, the LSQR error increases as well. Larger evaluation spheres have more nodes contained within them; the number of nodes increases in proportion to the cube of the radius. However, according to Equations 2.12 and 2.13, the number of equivalent sources (*NES*) only increases in proportion to the square of the radius. Therefore, large evaluation spheres use fewer equivalent sources per node. The black arrows in Figure 3.9 shows this ratio; for a 0.14-meter evaluation sphere radius, *NES* is almost equal to the number of nodes, whereas a 0.23-meter evaluation sphere has nearly twice as many nodes as equivalent sources. This results in a greater LSQR error as evaluation sphere radius increases (but recall from Figure 2.7 that overall auralization error decreases with radius).

Figure 3.11 includes four plots – one for each of four different evaluation sphere radii – indicating that as *SPWL* increases, the LSQR error decreases. However, at a quality level of approximately *SPWL* = 3.5, the addition of more equivalent sources appears to have a diminishing return. This observation agrees with the theory. For the largest evaluation sphere, the use of only 2.5 equivalent sources per acoustic wavelength is sufficient. Figure 3.12 includes four plots – one for each of four different octave bands – that show the LSQR error increases with frequency and evaluation sphere radius and decreases as *NES* increases. This trend is not obvious for the 125-Hz and 250-Hz octave bands (89 to 178 Hz, and 178 Hz to 356 Hz, respectively) since the LSQR error is comparable for all sphere sizes and *NES* values. The error for the 1000-Hz octave band (712 to 1428 Hz) is large when compared to the other three octave bands, but it the error plot is comparable to the theoretical LSQR error plot shown in Figure 2.5. The error discussed here only applies to the impulse response of the structure; the perceptual error associated with ESM due to other forcing functions is discussed in Chapter 5.

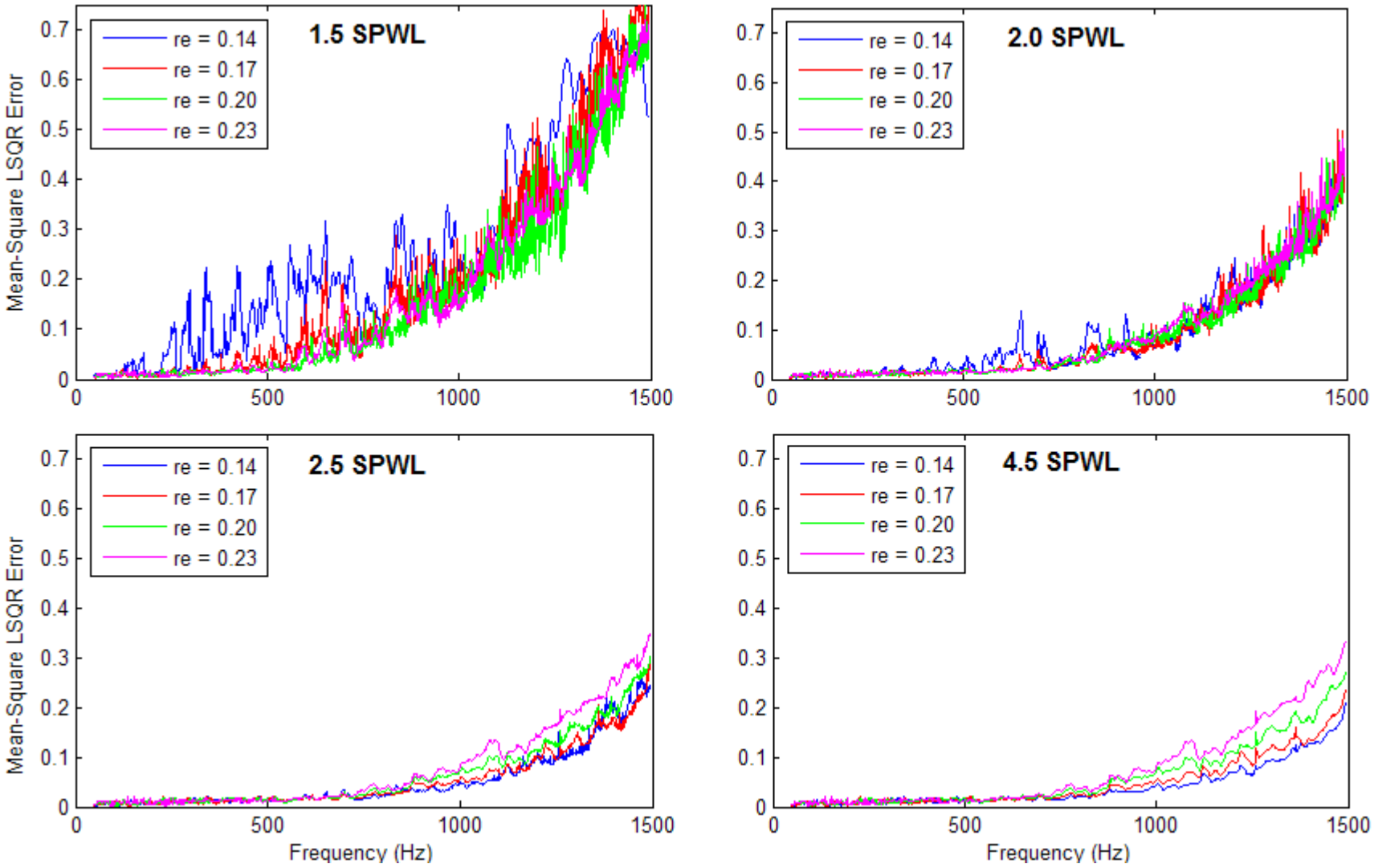

Figure 3.10. LSQR error as a function of frequency and evaluation sphere radius. Each plot is associated with a particular value of SPWL.

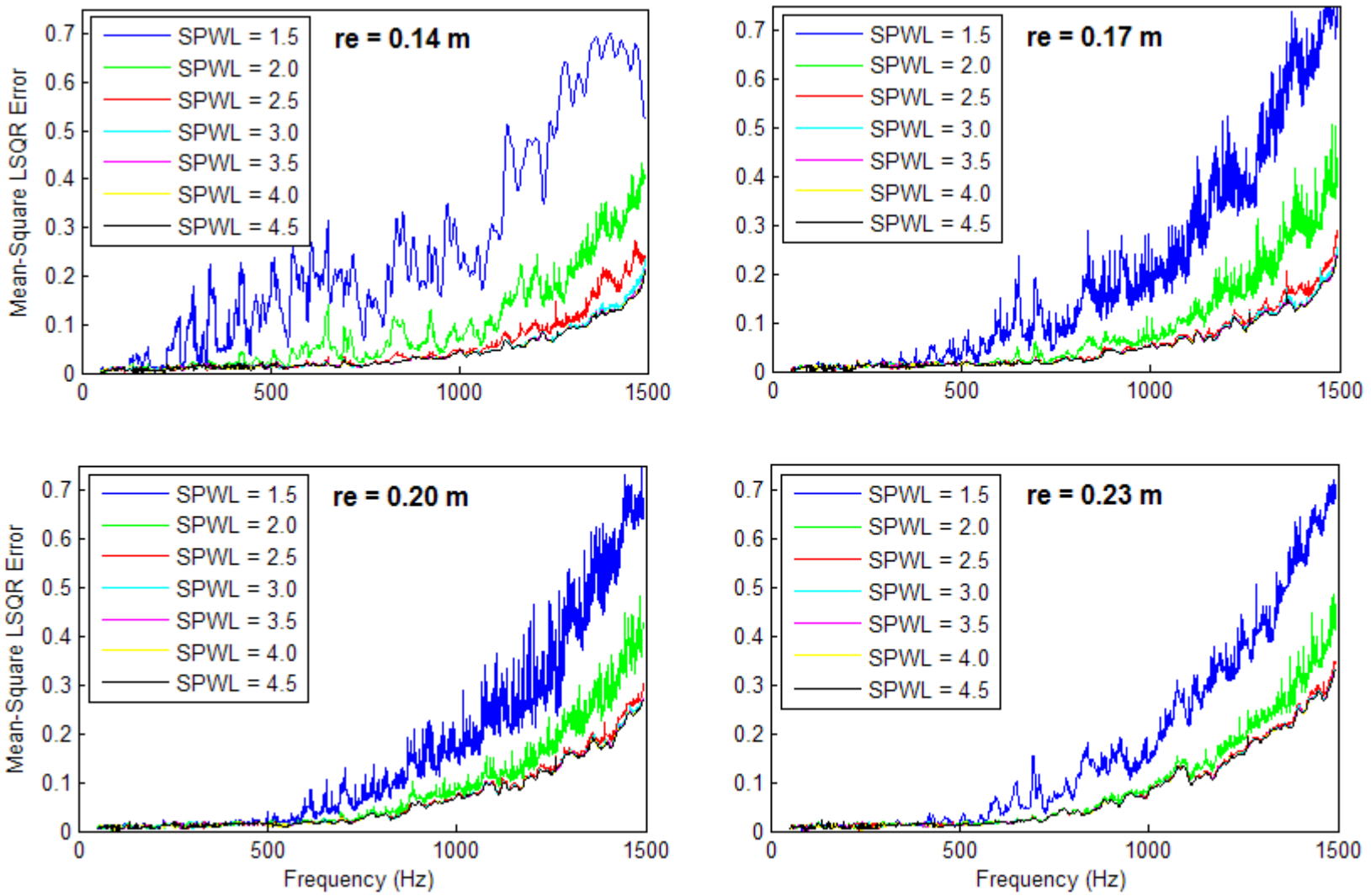

Figure 3.11. LSQR error as a function of frequency and number of equivalent sources per acoustic wavelength. Each plot is associated with a particular evaluation sphere radius.

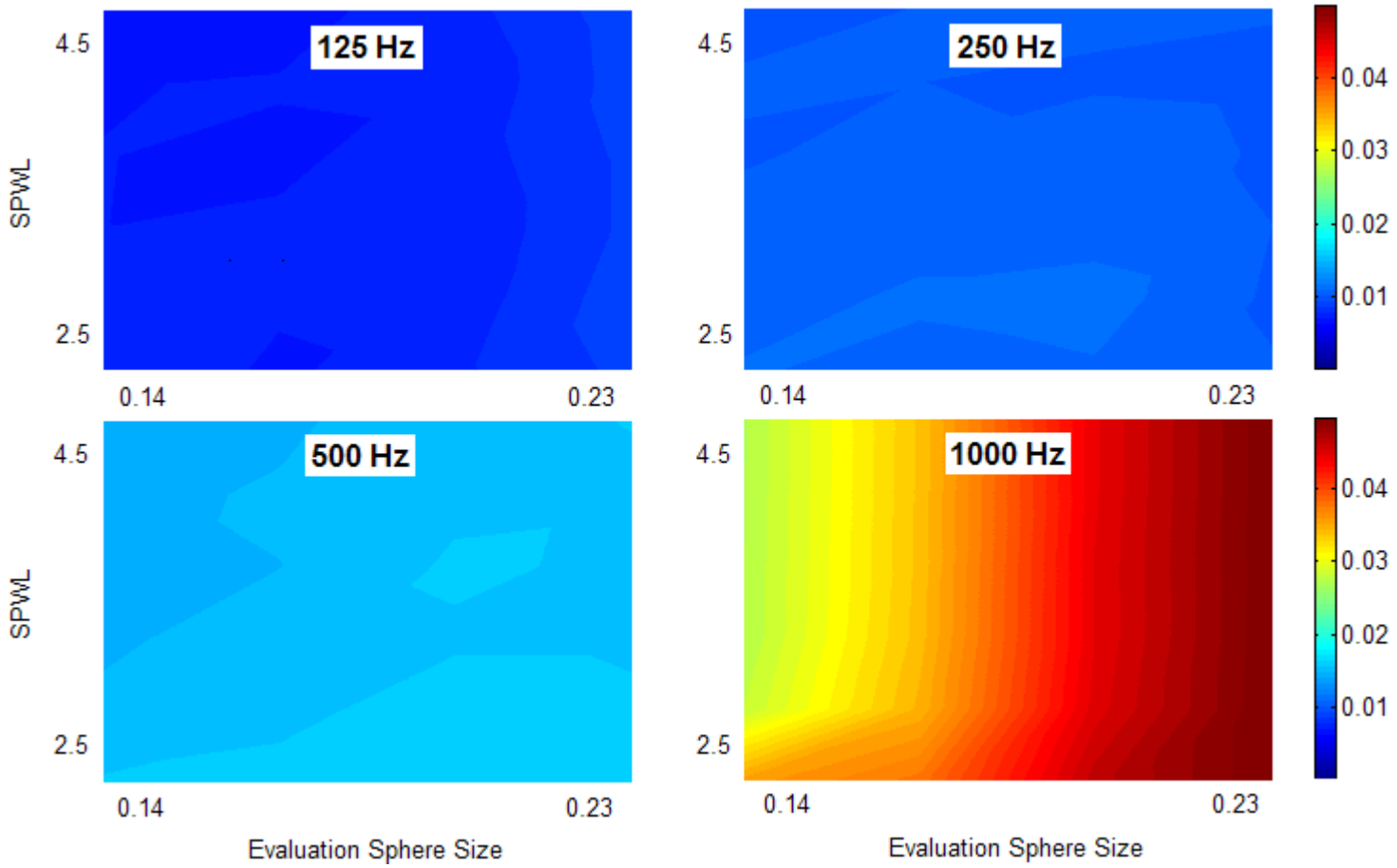

Figure 3.12. LSQR as a function of number of equivalent sources per acoustic wavelength and evaluation sphere radius. Each plot is associated with a particular octave band; the LSQR error was averaged across that octave band.

After calculating the frequency-domain equivalent source strengths **Q**, HRTFs were applied to the equivalent sources in order to calculate the left and right binaural signals (represented by b_L and b_R in the time domain, and B_L and B_R in the frequency domain). Since the head is at the geometric center of the evaluation sphere, all of the equivalent sources are at the same distance from the center of the head. Thus, the sound emanating from each equivalent source will reach the center of the head at the same time. Therefore, all of the sources can be summed in the frequency domain without being delayed in time. The left and right frequency-domain binaural signals can be calculated using

$$B_L(\omega) = \sum_{i=1}^{NES} Q_i(\omega) \cdot H_L(\theta_i - \theta_H, \phi_i - \phi_H, \omega) \tag{3.11}$$

$$B_R(\omega) = \sum_{i=1}^{NES} Q_i(\omega) \cdot H_R(\theta_i - \theta_H, \phi_i - \phi_H, \omega) \tag{3.12}$$

where Q_i is the frequency response function of the i-th equivalent source, and H_R and H_L are the left and right HRTFs, respectively.

The HRTFs are functions of acoustic frequency ω, head elevation angle ϕ_H, head azimuth angle θ_H, and the elevation and azimuth angles, ϕ_i and θ_i, between the *i*-th equivalent source and a reference direction. The reference direction is defined in the finite element model and points in the same direction as the positive x-axis. If the head was not centered in the evaluation sphere, the sources would need to be delayed by t_0, which is given by:

$$t_0 = \frac{d}{c} \tag{3.13}$$

where *d* is the distance between the source and the center of the head. The HRTFs can then be applied in the time domain and used to obtain the left and right time-domain binaural signals:

$$b_L(t) = \sum_{i=1}^{NES} q_i(t - t_0) \otimes HRIR_L(\theta_i - \theta_H, \phi_i - \phi_H, t) \tag{3.14}$$

$$b_R(t) = \sum_{i=1}^{NES} q_i(t - t_0) \otimes HRIR_R(\theta_i - \theta_H, \phi_i - \phi_H, t) \tag{3.15}$$

where $HRIR_L$ and $HRIR_R$ are the left and right head-related impulse responses (time domain equivalents of HRTFs), respectively, and q_i is the time response of the *i*-th equivalent source.

The HRTF set used in this thesis was measured on the Knowles Electronic Manikin for Acoustic Research (KEMAR) [33]. The HRTFs are defined over the full range of elevation and azimuth angles ($-90° \le \theta \le 90°$ and $-180° < \phi \le 180°$). Figure 3.13 shows the left and right HRTFs at two locations; the HRTF in the left plot corresponds to an azimuth of 45° and an elevation of 59.4° (above the head and slightly to the left), and the HRTF in the right plot corresponds to an azimuth of -45° and an elevation of 59.4° (above the head and slightly to the right). The azimuth angle closer to the right ear has a greater right-ear response, and the azimuth angle closer to the left ear has a greater left-ear response. The KEMAR set is symmetric with respect to the centerline of the body. Since the measurements were made on a manikin with carefully molded head and torso, this observation makes sense; the shape and size of the head and torso are primarily

responsible for shaping the HRTFs at low frequencies (at high frequencies, the pinnae geometry becomes important to the HRTF).

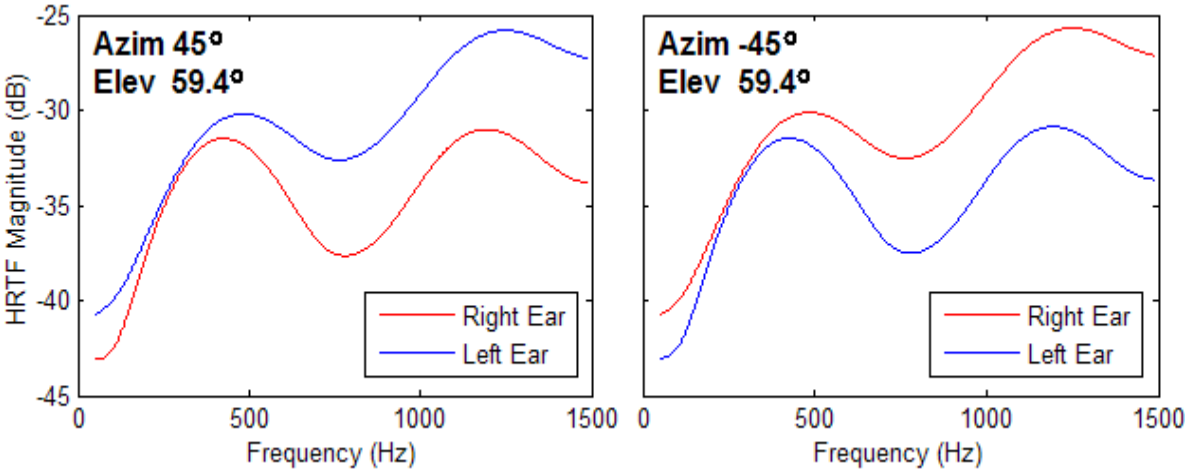

Figure 3.13. Left and right HRTFs for two azimuth angles located on opposite sides of the head. The left HRTF has a larger amplitude for azimuth angles between -180° and 0°; the right HRTF has a larger amplitude for azimuth angles between 0° and 180°. HRTFs are approximately symmetric about the head.

The frequency response function of the left and right binaural signals for four different head azimuth angles (θ_H = 0°, 45°, 90°, and 180°) is shown in Figure 3.14. The 0.23-meter sphere and 258 sources were used in the creation of the eight signals that are plotted. Low frequencies (between 50 and 600 Hz) and high frequencies (between 1200 and 1500 Hz) appear to dominate. The response is also unaffected by θ_H; although the individual sources should sound louder to the ear that is closer, the total binaural signal should sound just as loud at both ears. This means that regardless of their location on the source sphere, the equivalent sources have approximately equal strengths.

Figure 3.15 shows the time-domain left and right binaural signals. For θ_H = 0° – when the head is exactly facing the forcing location – the impulse response over the first 0.5 seconds after impact is nearly equal in magnitude and shape at both ears. For θ_H = 90° – when the right ear is facing the forcing location, the response just after impact is equal also nearly equal in magnitude at both ears. At the time of impact, the window is

excited, causing sound to be propagated towards the right ear. The structure's walls also begin to vibrate due to the impact and sound propagation; they are highly reflective, and the structure's acoustic response takes a long time – on the order of a few seconds – to die out after being excited. The multitude of reflections and sound emissions from the structure's MDF walls should reach the left and right ears with approximately equal amplitudes. The transient response of the structure is dominated by the first acoustic mode of the structure at 107 Hz.

The evaluation sphere radius and quantity of equivalent sources affect the frequency content of the binaural signal. Since a majority of the LSQR error occurs at high frequencies, the impulse responses will have different high-frequency content. High frequencies are attenuated very quickly, but are important to the sound of the initial impact. Figure 3.16 shows the percent difference between the frequency content of the best case and cases using smaller evaluation spheres and lower values of *SPWL*. The difference gets gradually larger as a smaller sphere and fewer equivalent sources are used. The psychoacoustic significance of this plot will be evaluated in Chapter 5.

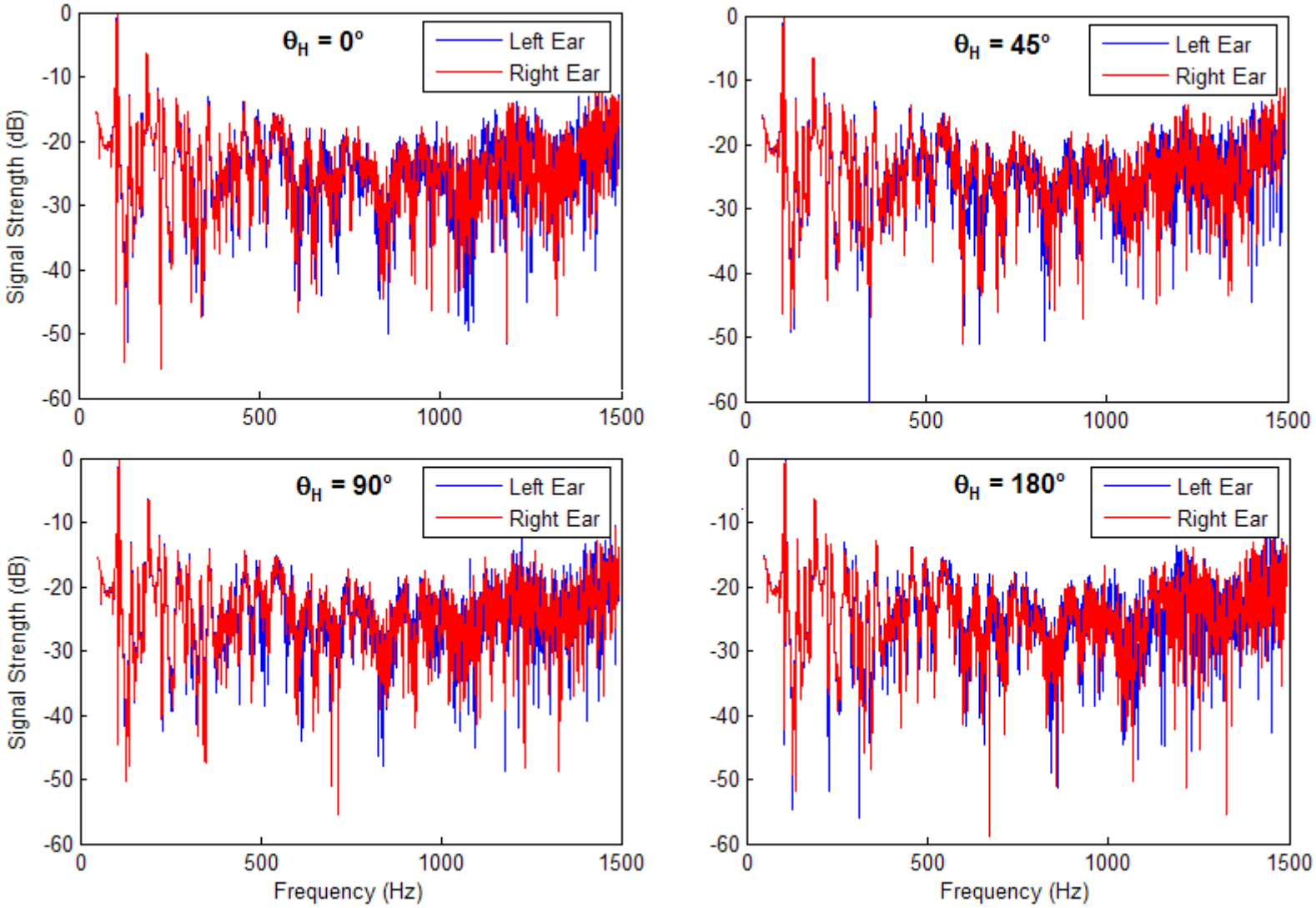

Figure 3.14. Frequency content of the left and right binaural signals for the largest evaluation sphere at the best quality. Four head angles are shown.

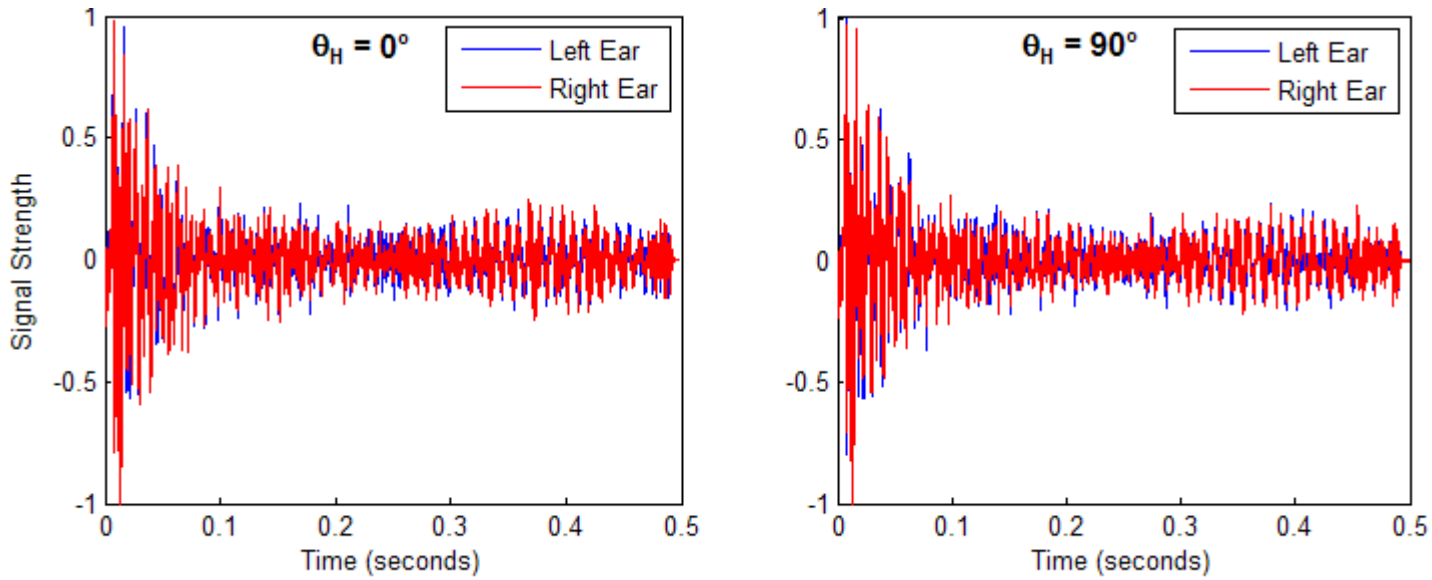

Figure 3.15. Time-domain left and right binaural signals for the largest evaluation sphere at the best quality. Two head angles are shown. The head angle does not significantly effect the relative strengths of the left and right signals.

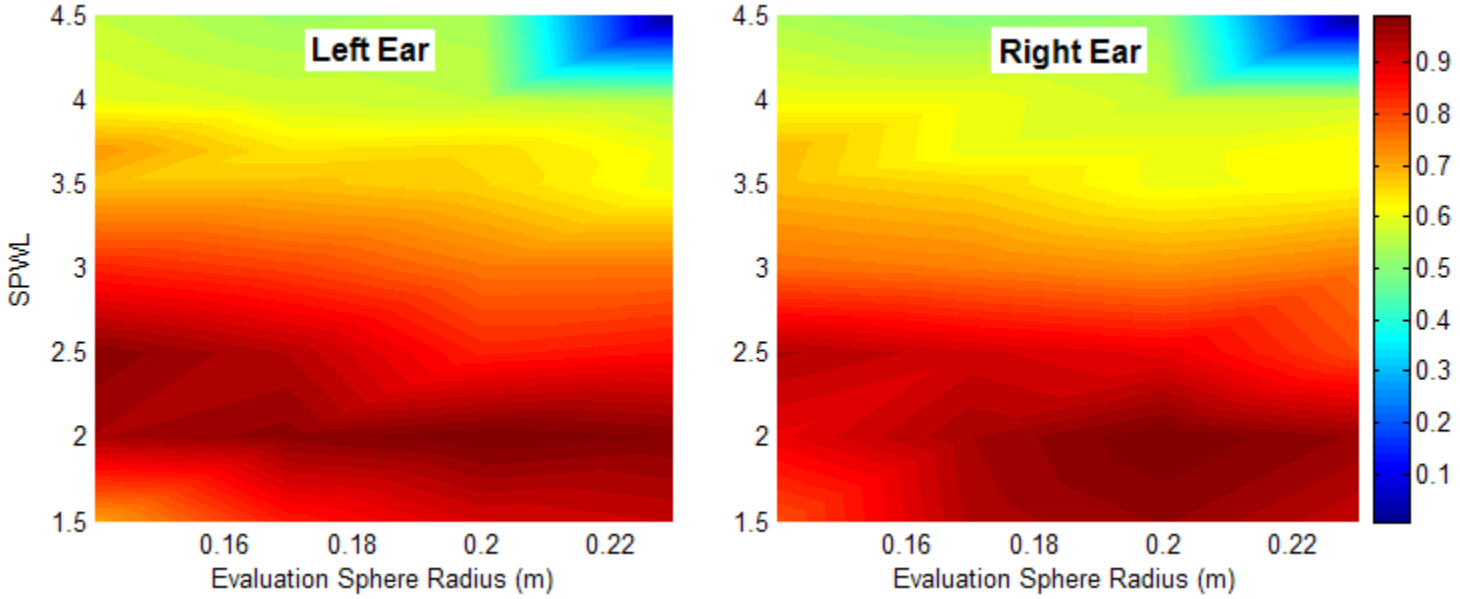

Figure 3.16. Percent difference between the frequency content of the best case (highest NES and largest evaluation sphere). The plots represent an average across the entire frequency band of 50 to 1500 Hz.

Chapter 4: Validation of the Numerical Model

This chapter will cover the construction of the structural acoustic test platform that was described in Chapter 3. The enclosure was built to validate the overall numerical model, which includes the finite element model, the equivalent source method (ESM), and the application of head-related transfer functions (HRTFs). Section 4.1 discusses the physical construction of the enclosure and how it differs from the model. Section 4.2 shows the measurements that were made on the enclosure and their relation to the finite element model output. Section 4.3 describes the insertion of the Knowles Electronic Manikin for Acoustic Research (KEMAR) into the enclosure and the measurement of the acoustic response at the ears to a structural force input. The response at the KEMAR's ears is compared to the hypothetical response calculated in Section 3.3.

4.1 Construction of the Structural Acoustic Test Platform

In order to validate the numerical model described in Chapter 3, a structural acoustic test platform was built in accordance with the design in Figure 3.4. The primary goal in design and construction was to hold a tight tolerance on the size of the enclosed acoustic space. See Figures 4.1 and 4.2 for photographs of the structure. Medium-density fiberboard (MDF) and Lexan were the primary building materials. MDF was chosen because of its low cost, wide availability, and durability; it is a conglomeration of tiny wood particles that have been drawn into fibers and pressed together with resin. It is representative of common building materials such as plywood and sheetrock. All of the MDF used for the enclosure was ¾" thick. It is very stable dimensionally, but the disadvantage to using MDF is that its modulus of elasticity can very greatly (2-5 GPa). For acoustic research purposes, MDF is still superior to other particle board and plywood. Lexan was chosen for the window material because of its high impact strength and machinability. Its modulus of elasticity is typically 900 MPa, but there is typically a large negative tolerance on the thickness. The ¼" sheet used in this test had some spots that thinner than 0.240". Lexan sheets with a nearly-exact uniform thickness are costly.

Figure 4.1. Structural acoustic test platform, complete but without door. The enclosure was constructed from six subassemblies that are all bolted together along their outside edges.

Figure 4.2. Structural acoustic test platform, complete with door. The door panel clamps into its slot from the outside, leaving a smooth interior wall.

The enclosure was constructed from six subassemblies that bolt together along their outside edges as shown in Figure 4.3. Wooden trim pieces were constructed from 1½" x 1½" x 8' sticks of lumber. The trim was fastened securely to the perimeter of each subassembly via wood screws. The large top and bottom panels were prevented from sagging by attaching additional trim pieces for support. Figure 4.3 shows the connection of three subassemblies (it is a detailed picture – taken from a different angle – of the top-right corner in Figure 4.2). The MDF panels were sized to overlap; when the bolts connecting the trim pieces on adjacent panels were tightened, the MDF panels were squeezed together to minimize the gaps along the seams on the inside. Figure 4.4 shows the interior of the corner in Figure 4.3; the seams are not obstructed by any trim pieces. The acoustic space was modeled as a five-sided extrusion with sharp corners in ABAQUS, so it is important that the enclosure was constructed with all of the bolted joints on the outside.

The windowed panel is the only side of the enclosure (aside from the door itself) that was constructed using double-ply MDF. The Lexan window is attached to the outer MDF panel and clamped down around its perimeter by ½"-thick aluminum bars. The inner MDF panel is attached to the outer panel by screws in the middle and bolts around the edges. The window is not attached to the inner panel. Figure 4.5 shows this connection in greater detail. The door panel is a double-ply piece of MDF that fits securely in a slot (shown in Figure 4.1) and gets clamped down (as shown in Figure 4.2). This design allows easy access to the acoustic space while maintaining a smooth interior wall.

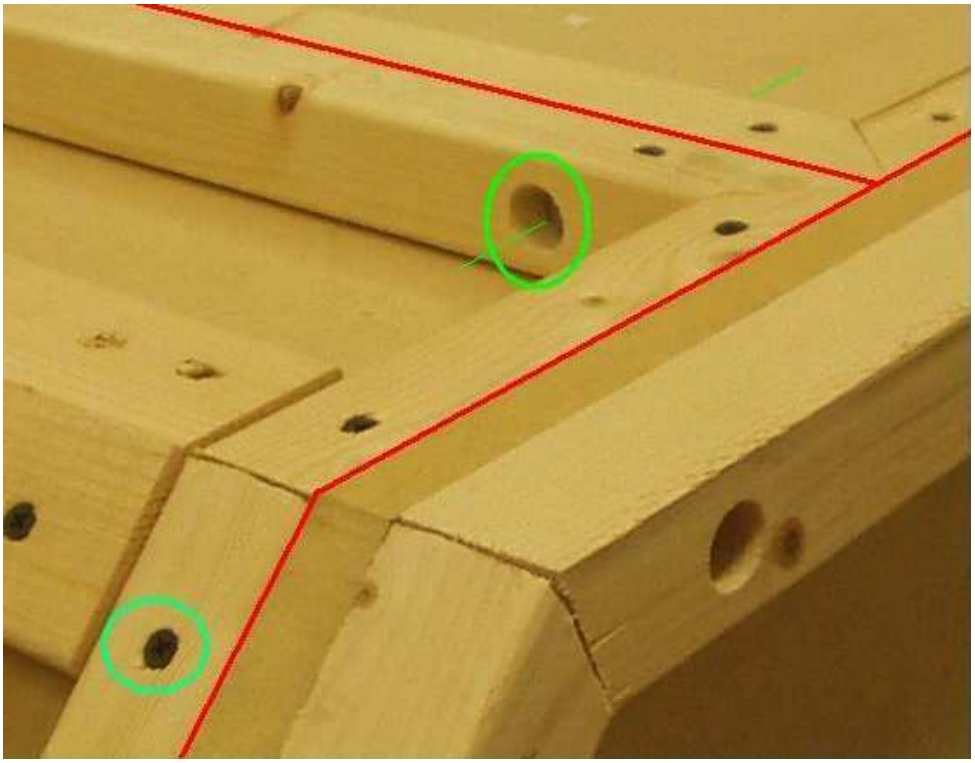

Figure 4.3. Connection of three subassemblies (distinguished by the red lines). The trim is attached to the MDF by screws (i.e. blue circle) and the subassemblies are bolted (i.e. green circle).

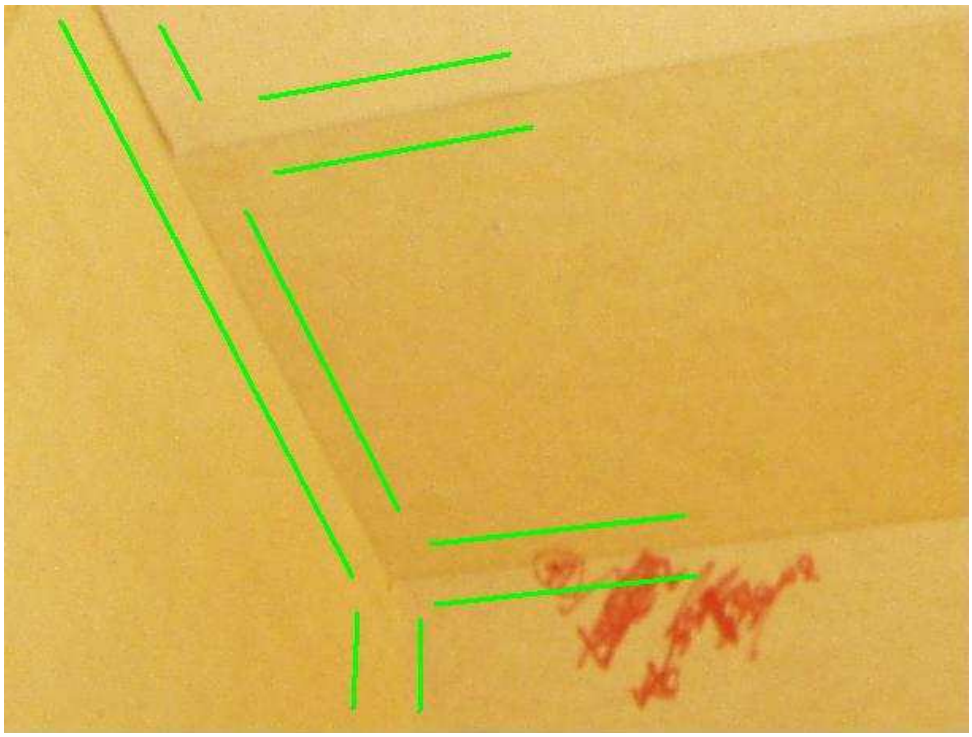

Figure 4.4. Interior of the connection shown in Figure 4.3. The seams (outlined in green for clarity) are not covered with any trim pieces, so the acoustic space has sharp corners as designed.

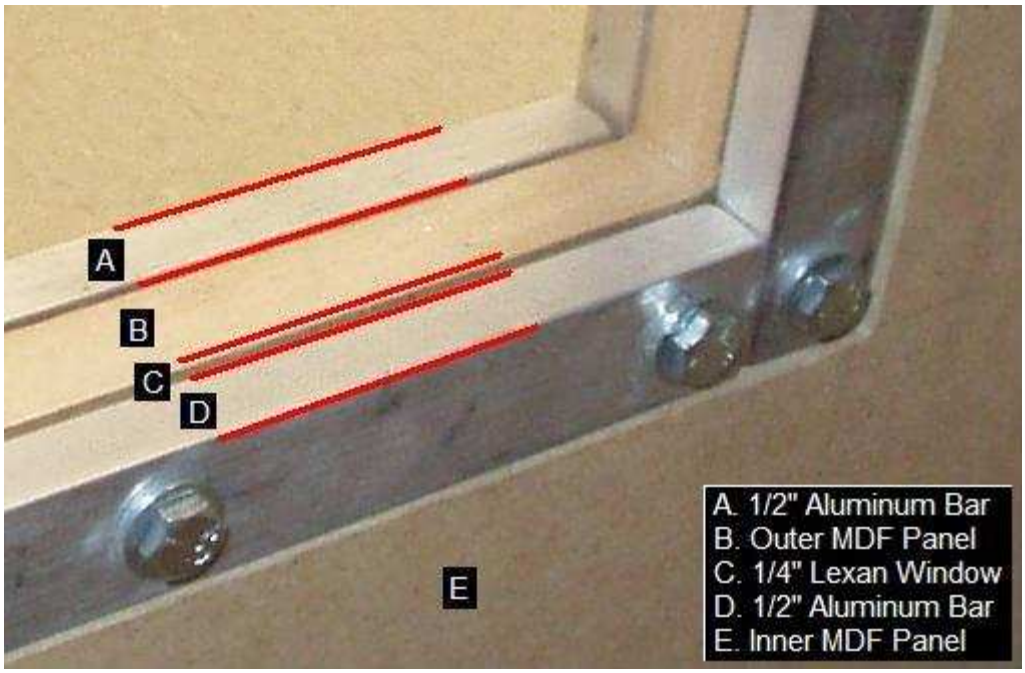

Figure 4.5. Attachment of the ¼"-thick Lexan window. The photograph was taken from the inside of the enclosure. The light conditions caused the window to appear deceptively thin.

The maximum dimensions of the interior of the enclosure are within ¼" of the 44½" x 49" x 69¼" acoustic space that was modeled in Chapter 3. The density of air, MDF, and Lexan used in the enclosure were measured before being entered into the model (see Table 3.1 for values). Although the modulus of elasticity and Poisson's ratio were not directly measured for the MDF and Lexan, the damping values were calculated from measured data. The bolted connections along the panel edges were approximated in ABAQUS by constraining the edges in six – three rotational and three translational – degrees of freedom. Modeling all of the trim pieces, screws, and bolts would be impractical; the lumber properties vary greatly from piece to piece, as does the clamping force due to the fasteners. These uncertainties are typical in the construction process.

4.2 Validation of the Finite Element Model

To check the finite element model for validity, the model output was compared with the measured frequency response functions (FRFs) of the structural acoustic test platform that is described in Section 4.1. To measure the FRFs, two types of forcing and two types of output were used. The acoustic space could be driven with a mid-range speaker in the corner of the enclosure near the door. The structure could be excited by a shaker that was attached to the Lexan window as shown in Figure 4.6. The shaker location is two inches over and two inches down from the top-right corner of the window. It was hung with two bungee cords so the body was free to move; the goal was to simulate the way an impulse force would excite the window. The stinger was attached to the window using superglue. A force transducer (PCB Piezotronics 208C01) was placed inline with the shaker's stinger to measure the input. Only one method of forcing was used at a time.

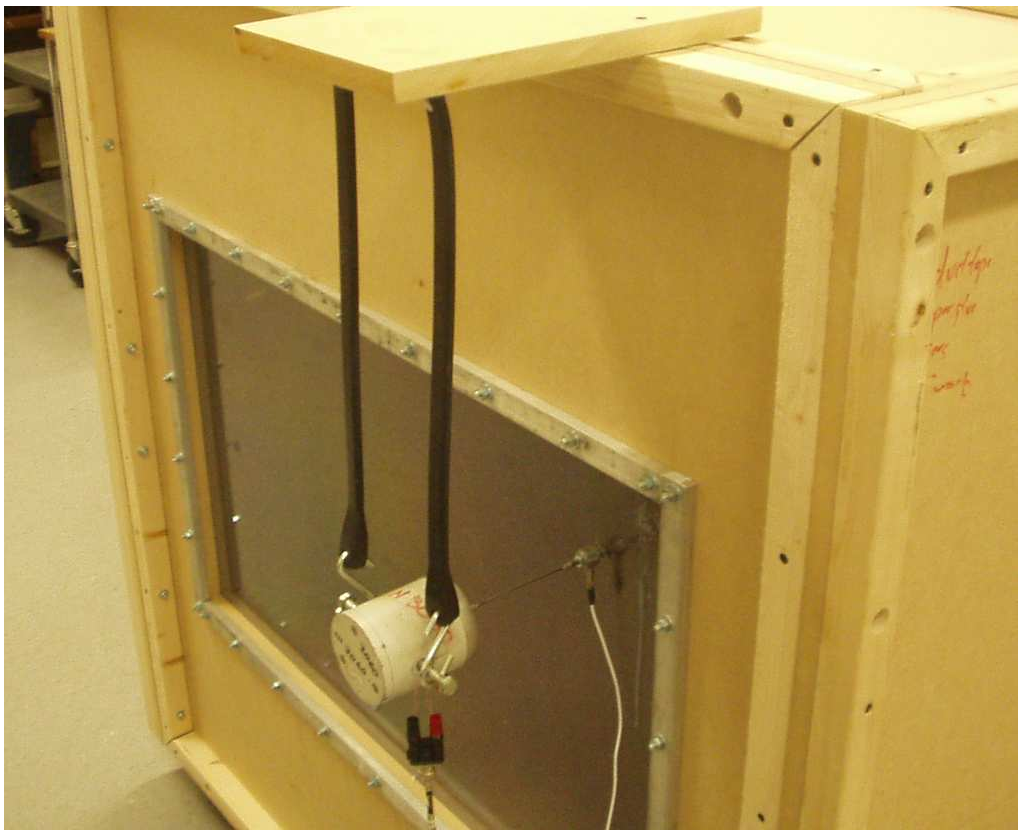

Figure 4.6. Attachment of the shaker to the Lexan window of the structural acoustic test platform.

The enclosure was instrumented with microphones and the MDF walls were instrumented with accelerometers. Both drive methods used a white-noise input that was high-pass filtered at 50 Hz and low-pass filtered at 2000 Hz. The noise was generated on a laptop computer and played over the headphone output jack. The signal was fed into a filter (Ithaco 4302 Dual 24dB/octave) and then through an amplifier (Yamaha HTR-5440 receiver). The microphone, accelerometer, and force transducer outputs were passed through a signal conditioner (PCB Piezotronics 433A with Patch Panel 070A17) and fed into a National Instruments SCXI-1001 data acquisition chassis. The sampling rate was 8192 Hz, and ten four-second samples were taken for each transfer function. Four types of transfer functions were measured:

- Microphone output per speaker input
- Accelerometer output per speaker input
- Microphone output per shaker input
- Accelerometer output per shaker ouput

The first observation was that the microphone FRFs were remarkably similar for both speaker and shaker forcing. FRFs for both types of forcing contained acoustic modes and structural modes. As expected, the shaker excites the structure and the acoustic space. In addition, the sound emanating from the speaker causes the structural walls to vibrate. The accelerometer FRFs indicate this phenomenon as well; all of them contain both acoustic and structural modes. The magnitudes of the wall vibrations are approximately the same for shaker and speaker input (with the exception of measurement points close to the shaker). Since the shaker and speaker do not excite the structural acoustic test platform differently, only the shaker was used for the remainder of this analysis.

The comparison between the finite element model and the enclosure's actual response was performed in two stages. Initially, the frequency range between 50 Hz and 250 Hz was investigated. The six most dominant acoustic modes were measured, and the material properties were altered slightly to make their corresponding natural frequencies match as closely as possible. As discussed in Section 3.2, there is a degree of uncertainty associated with the densities, Young's moduli, and damping coefficients in the finite

element model. Following adjustment of the model, the full frequency range – 50 Hz to 1500 Hz – was investigated. The damping coefficients of the MDF were modified until the modal density and peak heights in the magnitude plot of the model's frequency response matched the magnitude plot of the measured FRFs.

To measure the acoustic modes between 50 and 250 Hz, the planar microphone array shown in Figure 4.7 was constructed from PVC pipe and fittings. The perimeter of the array has the same shape as the cross section of the structural acoustic test platform; it is placed at different depths of interest in the enclosure. The pipes can be reconfigured to produce even more measurement locations. Only 30 microphones were used, but 472 FRFs were measured by using the array in eight locations and two configurations. The measurements showed that the acoustic natural frequencies were roughly 2-6% higher than the model predicted. The density of air was lowered to 1.15 kg/m^3, which according to the ideal gas law is a good estimate for the air in Blacksburg, Virginia. Damping was added to the model as well, since the original peaks were too large; the first model did not include any damping. All other parameters were unchanged. Figure 4.8 shows the final comparison between the finite element model output and the measured "microphone output per shaker input" between 50 and 250 Hz. The six most dominant acoustic natural frequencies (107, 143, 173, 191, 218, and 235 Hz) are highlighted with magenta circles. Both curves use data from Location 2, which is indicated in Figure 4.9. For more information about how the planar microphone array was used to collect and evaluate the structural-acoustic data, refer to Appendix A.

Figure 4.7. Planar microphone array. This setup allows for the rapid measurement of "microphone output per shaker input" FRFs at all locations inside the enclosure.

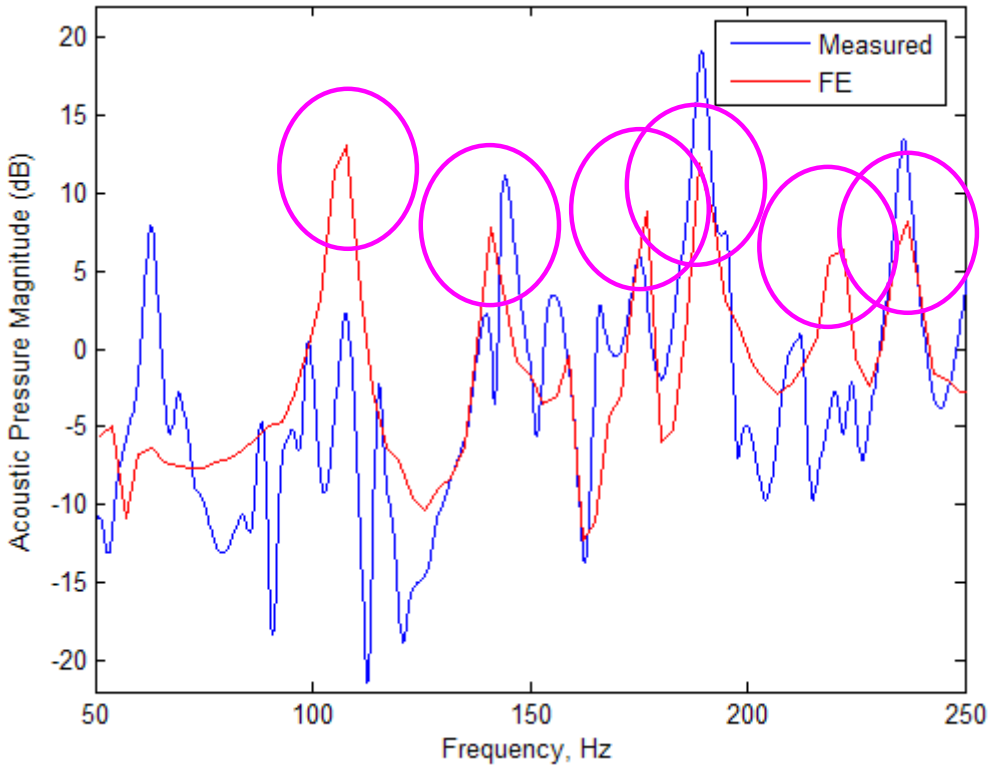

Figure 4.8. Final comparison between the finite element model output and the measured "microphone output per shaker input" between 50 and 250 Hz. Five of the six most dominant calculated acoustic modes appear to be present in the measured enclosure FRF.

After a satisfactory match between the model and the measured FRFs was obtained between 50 and 250 Hz, the solution was calculated for the full frequency range of 50 to 1500 Hz. Since ABAQUS calculated the response of the enclosure to a sinusoidal force input at each frequency, and the impulse response is of interest, each point in the modeled FRF was divided by $\omega/((j\omega)^2+\omega^2)$ to obtain the impulse response FRF. The subsequent output showed just as many modes as the measured FRFs, but the peaks were significantly higher than the measured peaks. This result indicated that the damping coefficients needed to be raised in the model. However, the addition of more damping would the modes to have lower peaks and become indistinguishable in the magnitude plot. Suitable α and β values were chosen to strike a balance between modal peak height and modal density. The final model parameters are those listed in Table 3.1. The output was calculated in ABAQUS using the final parameters; Figure 4.10 shows the comparison between the model and the measured FRFs at Location 1 and Location 2 (shown in Figure 4.9).

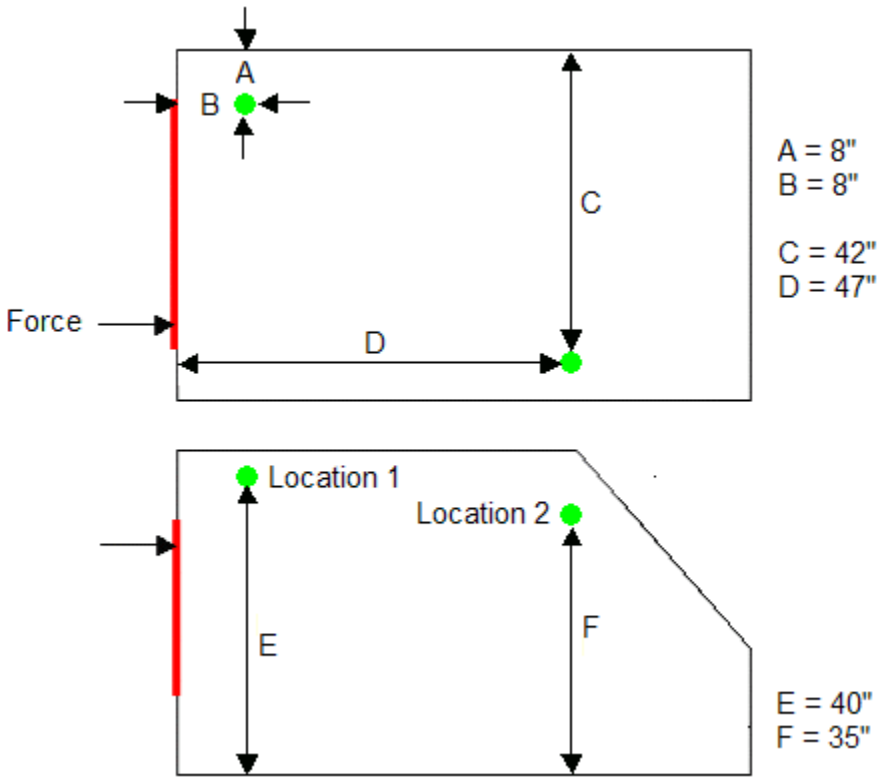

Figure 4.9. Locations used in the comparison of the finite element model output to the measured FRFs.

The correlation between the finite element model output and the measured FRF at was 36.1% at Location 1 and 35.3% at Location 2. From Figure 4.10, it is apparent that a difference in the FRFs exists between 800 and 1500 Hz. The model response contains many modes, and thus an extremely high modal overlap factor is present; few modes stand out and dominate the response. The physical enclosure has many more modes between 800 and 1500 Hz that arise from elements that were not modeled such as the door and stiffening ribs that are attached to the backside of the panels. The model is symmetric and the structure is not (due to imperfect construction), so several instances of mode splitting appear in the measured response. Although, numerically, the correlation coefficient is low, the binaural signals that were created from the finite element model output sounded very similar to the physical excitation of the enclosure. The numerically-large discrepancies occur as a result of small deviations in the material properties and boundary conditions from their modeled values. Few of the high-frequency modes overlap, but the two very-different frequency responses can occur in two very similar instances of the same structure [20].

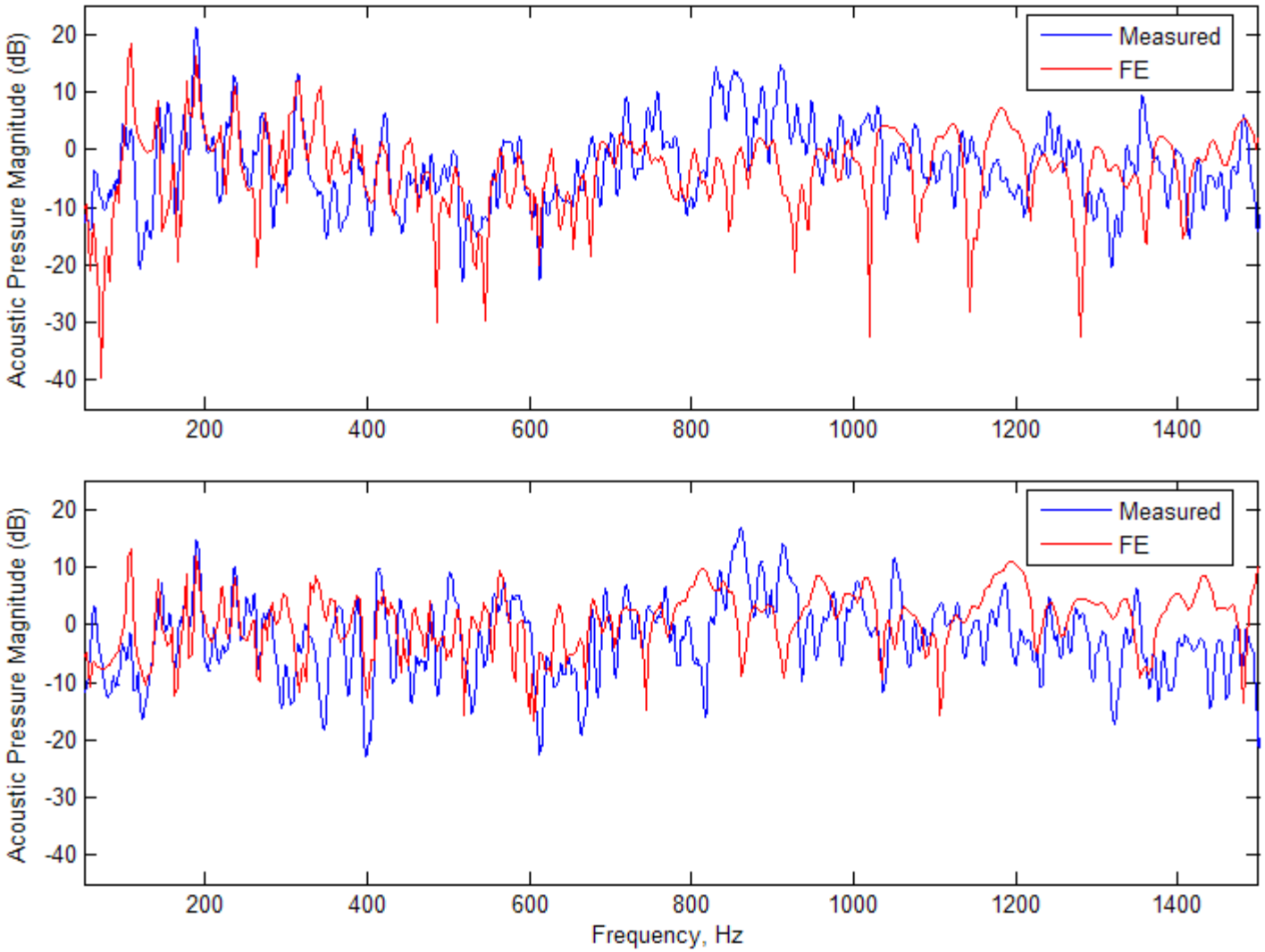

Figure 4.10. Comparison between the finite element model output and the measured FRFs of the enclosure. Location 1 (top) and Location 2 (bottom).

4.3 Manikin Measurements

The binaural acoustic response at the ears of the Knowles Electronic Manikin for Acoustics Research (KEMAR) was calculated using a finite element model and ESM in Section 3.3. After obtaining a satisfactory match between the predicted response from the finite element model and the measured response of the structural acoustic test platform, a KEMAR was placed in the enclosure as shown in Figures 4.11 and 4.12. The manikin was oriented such that it faced the shaker, which was coincidentally at nose-level. The response was measured at both ears using Brüel and Kjær microphones. Their output was passed through a Brüel and Kjær signal conditioner and into a National Instruments SCXI-1001 data acquisition chassis. For each signal, ten four-seconds samples were recorded at 8192 Hz. The transfer functions between the left and right ear signals and the shaker input (taken from a force transducer) were measured.

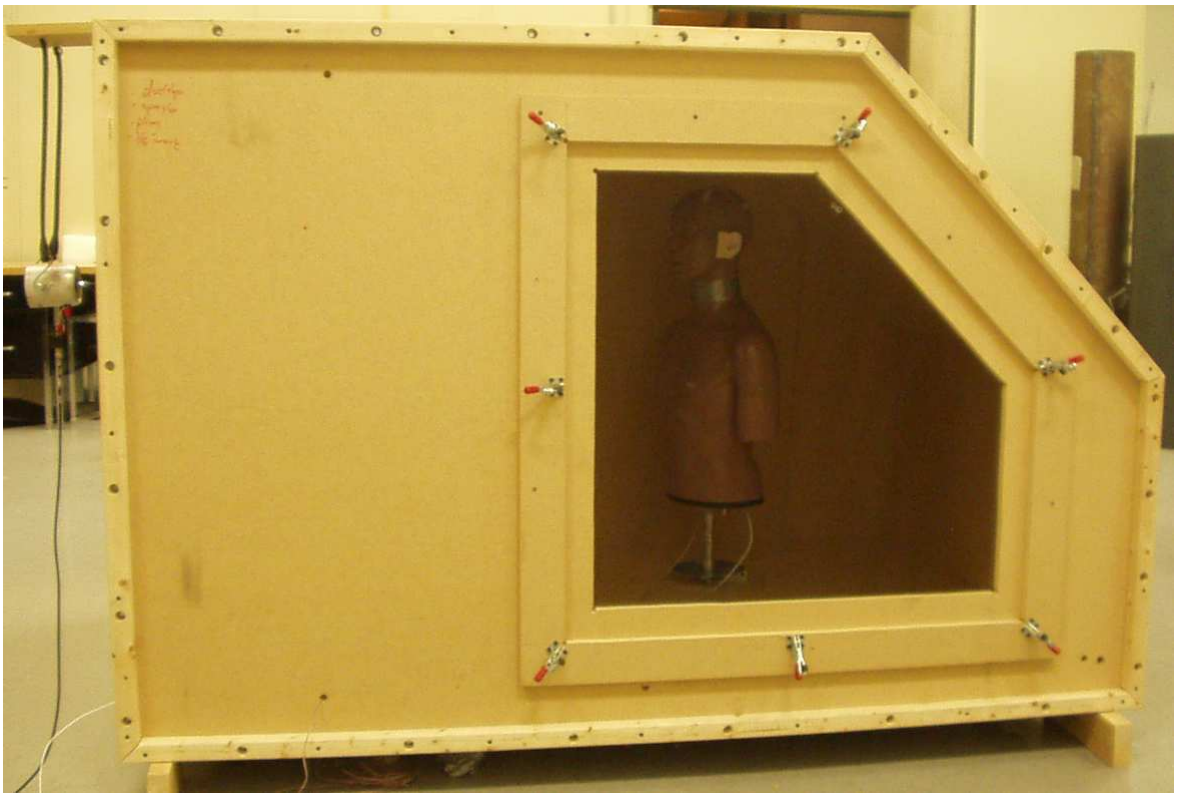

Figure 4.11. KEMAR inside the structural acoustic test platform. The response at the ears was measured and compared to the calculated response shown in Section 3.3.

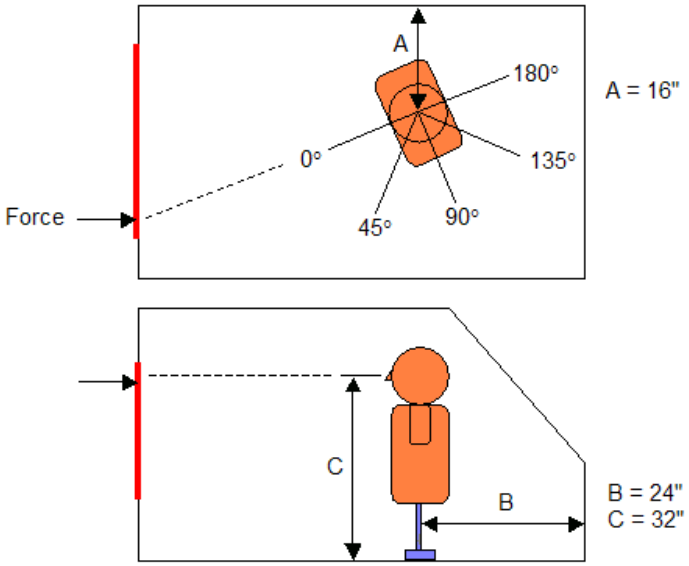

Figure 4.12. Position of the KEMAR in the enclosure.

Binaural signals were measured for five head azimuth angles (θ_H = 0°, 45°, 90°, 135°, and 180°) at head elevation angle ϕ_H = 0°. For the first and last angles, the left and right ears are at the same distance and angle from the shaker, but for the other three cases, the right ear is spatially closer and aimed closer to the shaker. In theory, the right and left ear impulse response signals should be nearly identical for θ_H = 0° and θ_H = 180°. The initial impact should sound the same at both ears, and the reflections will have similar frequency content. Since the enclosure is not symmetric about the centerline of the head, differences between the left and right signals surface. For the θ_H = 45°, θ_H = 90°, and θ_H = 135° cases, the initial impact should reach the right ear first. In addition, the magnitude of the impact should be greater at the right ear. Reflections will come from all directions, so the frequency content of the left and right signals should be similar, just as in the θ_H = 0° and θ_H = 180° cases. The intensity and arrival time of each reflection differs for each ear. The difference between the left and right signals should be greatest for θ_H = 90°, and the difference should increase with frequency.

Figures 4.13 and 4.14 show a comparison between the measured and calculated binaural ear responses for four head angles (θ_H = 0°, 45°, 90°, and 180°). The calculated binaural signals were generated by applying ESM and HRTFs to the finite element model

output as outlined in Section 3.3. The acoustic response of the enclosure to a sinusoidal input was determined in ABAQUS, and the output was divided by $\omega/((j\omega)^2+\omega^2)$ at each frequency to obtain the impulse response. ESM was applied for four different evaluation sphere radii (ranging from 0.14 to 0.23 meters) and eight different quality levels (ranging from 1.5 to 4.5 equivalent sources per acoustic wavelength). The KEMAR HRTF set was applied to the equivalent source sets to produce the left and right binaural responses. The modeled responses shown in Figures 4.13 and 4.14 use the largest evaluation sphere and the greatest number of equivalent sources (258). The effect of using fewer sources and smaller evaluation spheres will be investigated psychoacoustically in Chapter 5.

Figure 3.13 shows azimuth angles $\theta_H = 0°$ and $\theta_H = 180°$; the manikin is facing directly towards the shaker and directly away from the shaker, and as predicted, the measured binaural signals are approximately equal in magnitude and frequency content. The interaural time delays, interaural intensity delays, and pinnae effects are the same at both ears since the manikin is symmetric. However, there are more clearly defined modes between 800 and 1500 Hz at the left ear in the first case and the right ear in the second case. This indicates that the doored MDF panel is has more vibration modes in that frequency range than the non-doored panel. This phenomenon would not be apparent in the modeled response because the door was not modeled; the panel was treated as though it were a continuous sheet of MDF.

Figure 3.14 shows azimuth angles $\theta_H = 45°$ and $\theta_H = 90°$; the manikin is oriented with its right ear pointing closer to the shaker. In the second case, the manikin is facing the door. The modeled responses are the approximately the same at both ears for these head angles, but the measured responses differ. The right ear responses are similar in magnitude to the left ear responses between 50 and 600 Hz. From 600 to 1200 Hz, the average response is several decibels greater at the right ear. At higher frequencies, the left and right binaural signals become equal in magnitude again, indicating that the response is dominated by reflection-induced vibration of the MDF walls in this range.

The numerical correlation between the measured responses and the calculated signals is low. However, upon initial listening, the sounds of the modeled and actual binaural signals are very similar. Analysis of the binaural signals in the time domain will be performed numerically and subjectively in Chapter 5.

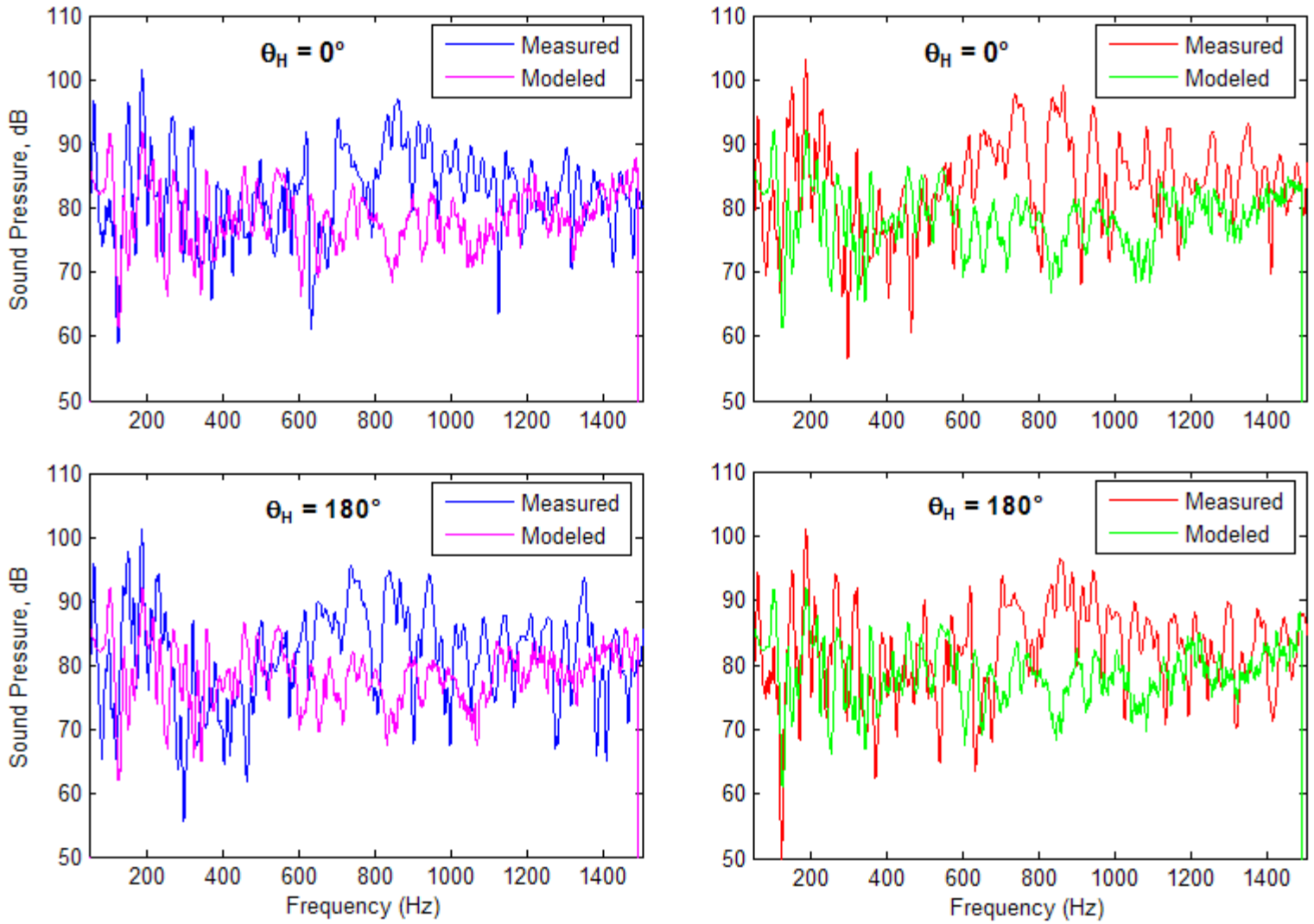

Figure 4.13. Comparison between the measured and calculated responses at the left ear (two left plots) and right ear (two right plots). For these two angles, the left and right signals should have approximately equal magnitudes.

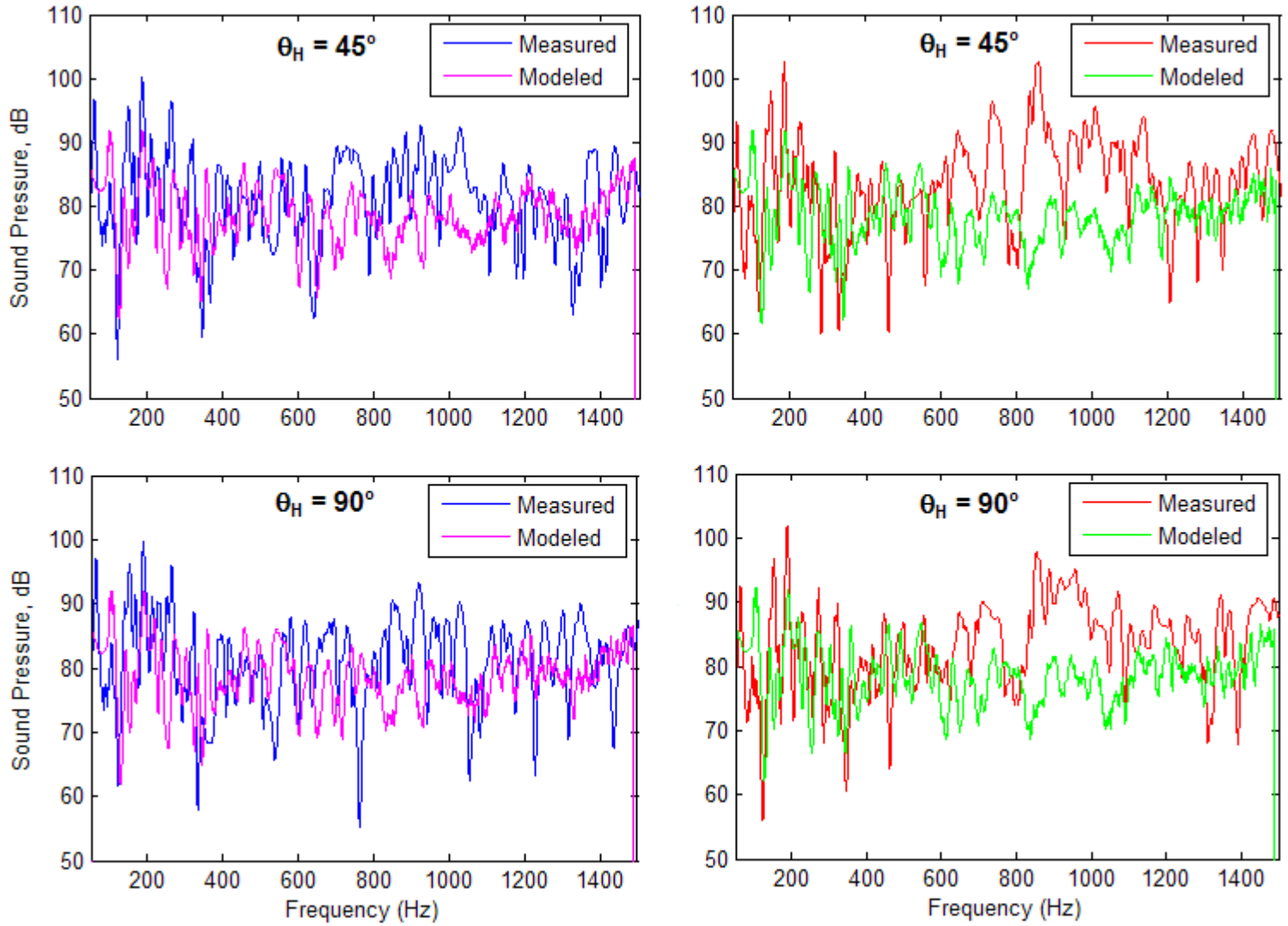

Figure 4.14. Comparison between the measured and calculated responses at the left ear (two left plots) and right ear (two right plots). For these two angles, the right signal should have a greater magnitude that the left signal.

Chapter 5: Psychoacoustic Testing

Chapters 3 and 4 discussed the numerical associated with the virtual acoustic prototyping process. In the same way the finite element model – both input and output – does not perfectly represent the physics of the enclosure that was constructed, the equivalent source method (ESM) cannot exactly reproduce the pressure field that was calculated in ABAQUS. Head-related transfer functions (HRTFs) are only discrete approximations of a typical head, so their application only closely estimates the binaural signals that a person would hear. This chapter discusses the psychoacoustic testing that was performed in order to relate numerical error to perceptual differences in the comparison of two signals. Section 5.1 explains the reason for subjective testing, and Section 5.2 describes the interface and test procedure. The results of the subjective testing, including the allowable error level for ESM, are discussed in Section 5.3. The signals used in the subjective testing were preprocessed; a real-time head-tracked configuration was not used because the only available real-time auralization system had not been fully debugged at the time of testing.

5.1 Reason for Psychoacoustic Testing

As discussed in Section 2.3, an auralization can be configured to produce very low error level if sufficient computational power is available. By using careful finite element modeling (FEM) techniques, larger evaluation sphere, a greater number of equivalent sources, and densely-measured HRTFs, a binaural signal can be produced over a set of headphones that will give a listener the impression that he is actually in the modeled structure. However, numerical difference between two signals does not necessarily cause a perceptual difference between two sounds. Figure 5.1 illustrates this point; the plot on the left contains a decaying sine wave (red) and a copy of this signal with a slight delay in time (blue). It is obvious from the substantial amount of white space between the two signals that the numerical difference is relatively large. However, if the two signals were played sequentially over a set of headphones, a listener would not

be able to determine which one contains the delay. The plot on the right contains a decaying sine wave (red) and a slightly noisy copy of this signal (blue). The numerical difference between these two signals is very small. If the signals were played sequentially over a set of headphones, a listener would be able to pick out the noisy signal very easily.

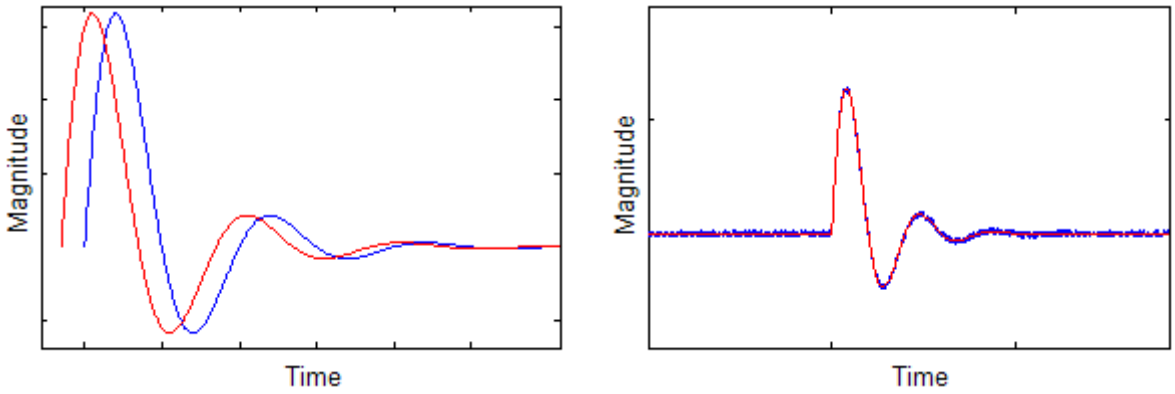

Figure 5.1. Two signal comparisons. There is a large numerical difference (and almost no perceptual difference) between the two signals in the left graph, and a large perceptual difference (and a small numerical difference) between the two signals in the right graph.

The goal of the psychoacoustic testing described in this section is to determine the evaluation sphere radius and the number of equivalent sources per acoustic wavelength that are needed to create an accurate auralization. Accuracy is determined by whether or not a typical listener can hear a difference between two binaural signals. One reference signal is created using the greatest number of equivalent sources and the largest evaluation sphere. The other signal uses a smaller number of sources and an equal-sized or smaller evaluation sphere. The evaluation sphere size and number of sources can be chosen to provide an acceptable level of error without exceeding the maximum available computational power.

5.2 Interface and Test Procedure

The psychoacoustic testing in this thesis consists of a 45-minute sound comparison exercise. Pairs of sounds created from the binaural output of the virtual acoustic prototyping process were played over a set of headphones and the test subject specified whether of not he could hear a difference. Each subject performed 204 comparisons; three different sound types, four evaluation volume sizes, and eight equivalent source quantities were used.

According to Figure 2.7, the lowest numerical error in the binaural signal occurs for a large evaluation sphere and many equivalent sources. Figure 5.2 shows the 32 error levels that were used in the test – with the lowest error being in the top-right corner. The magenta and blue circles indicate the cases for evaluation sphere radii 0.14, 0.17, 0.20, and .23 meters using 1.5, 2.0, 2.5, 3.0, 3.5, 3.7, 4.0, and 4.5 equivalent sources per acoustic wavelength (*SPWL*, refer to Section 2.3). All but nine comparisons were between a particular noise type at the lowest error level, indicated by the magenta circle, and the same noise type at one of the other error levels. Nine tests of the reference case versus itself – three for each noise type – were included to determine the trustworthiness of the subject's responses. The red circles in Figure 5.2 show the number of evaluation points that were used to calculate the equivalent sources, and the green circles show the total number of nodes in the evaluation sphere. The test cases were chosen such that the number of equivalent sources did not meet or exceed either of these quantities. A finer mesh would allow more equivalent sources, but there is no evidence that more sources would significantly reduce the numerical error.

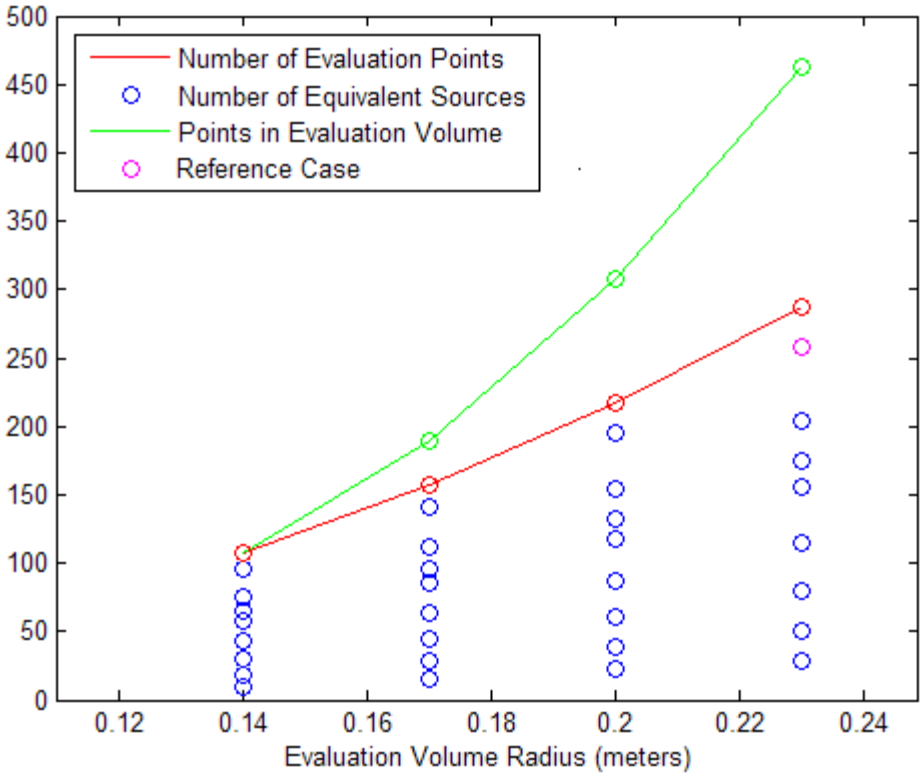

Figure 5.2. Subjective test cases. Binaural signals at 32 error levels (magenta and blue circles) were compared to the signal that uses the largest evaluation sphere and the most sources (magenta circle).

The three noise types used in testing were impulsive, tonal (engine noise), and broadband. The finite element model output is the response of the acoustic space to an impulse force on the window. Thus, the binaural signal that results from the application of ESM and HRTFs is a rapidly-decaying sine wave; the sound is as if a window was being tapped or hammered. To obtain the structural response to a different forcing function, then that forcing function must be convolved with the impulse response. In this testing, the binaural impulse response was convolved with time signals consisting of:

- a series of seven harmonics (200, 400, 600, 800, 1000, 1200, and 1400 Hz). Each harmonic had a lower amplitude than the one before it. The resulting time signal sounds like the horn of a car.
- white noise with frequency content between 50 and 1500 Hz. The resulting time signal sounds like the interior of a car on the interstate or the interior of an aircraft cabin.

- a Hanning pulse. This convolution gave the impulsive sounds that were played during testing; A pure impulse has zero width (which does not exist in practice). The Hanning pulse has a width on the order of one millisecond, and the resulting time signal sounds like a realistic structural impact.
- engine noise. This sound consists of eighteen harmonics with amplitudes that increases slightly with time. The sixth, twelfth, and eighteenth harmonics are dominant. The resulting time signal sounds like a revving V-6 engine.

The spectograms of these signals are shown in Figure 5.3. Only the Hanning pulse, white noise, and engine noise were played to the test subjects (the reason for this is that ESM does not appear suitable for use with a purely tonal signal; this will be explained in greater detail in Section 5.3). Each of these three sounds was convolved with the 32 binaural impulse responses that correspond to the error levels. A total of 96 sounds were created and used in the psychoacoustic testing.

Each noise type was used in 68 different comparisons; 17 tests of each noise type correspond to one of four evaluation sphere sizes. Table 5.1 shows the matrix of tests. Matrix of test cases. S1, S2, S3, and S4 are the evaluation sphere sizes from smallest to largest, and N1, N2, and N3 are the noise types (impulsive, tonal, and broadband, respectively). The sphere size and noise types are listed at the top of each column, and the quantities of equivalent sources – which will be termed "quality" from now on – are listed in the rows (with "1" being the lowest quality of 1.5 *SPWL* and "8" being the highest quality of 4.5 *SPWL*). The first seven rows indicate comparisons between the reference sound and a less accurate sound. The lowest-error sounds are shown with red text, and correspond to the largest evaluation sphere and the highest quality. Rows eight through ten all use the highest quantity of equivalent sources (4.5 *SPWL*), but varying sphere sizes. The nine numbers shown in red signify the repeat cases, and are intended to test the credibility of the subject's answers. The final seven rows are labeled "It" and indicate the iterative tests. For each sphere size and noise type, seven iterative comparisons are made. The first iterative comparison uses quality "5", and next six tests use a quality level that depends on the subject's prior iterative comparison answers for that particular sphere size and noise type. If the subject indicates that he can hear a

difference in the iterative comparison, then the next iterative comparison will use a higher quality; if the subject indicates that he cannot hear a difference twice in a row (for the same pair of sounds), then the next comparison will use a lower quality.

Table 5.1. Matrix of test cases. S1, S2, S3, and S4 are the sphere sizes from smallest to largest, and N1, N2, and N3 are the noise types (impulsive, tonal, and broadband, respectively).

S1N1	S2N1	S3N1	S4N1	S1N2	S2N2	S3N2	S4N2	S1N3	S2N3	S3N3	S4N3
1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8
It	It	It	It	It	It	It	It	It	It	It	It
It	It	It	It	It	It	It	It	It	It	It	It
It	It	It	It	It	It	It	It	It	It	It	It
It	It	It	It	It	It	It	It	It	It	It	It
It	It	It	It	It	It	It	It	It	It	It	It
It	It	It	It	It	It	It	It	It	It	It	It
It	It	It	It	It	It	It	It	It	It	It	It

At the beginning of each subject's test, a 204-by-2 "order matrix" was created to store the sequence of the test cases shown in Table 5.1. The elements of Table 5.1 were chosen in a random order and their row and column numbers were placed in the 2-element rows of the order matrix. The noise types, evaluation sphere sizes, and quality levels were mixed so the subject could not acclimate himself to a particular sound. Beginning at the top of the order matrix, the pairs of sounds were played to the subject, and after each pair was played, he was asked to indicate whether he could hear a difference. The test cases and the subject's responses were stored in a data file.

Figure 5.4 shows the National Instruments' Labview interface that was used in the psychoacoustic testing. Once the subject agrees to participate in the comparison exercise, he is directed to the interface. After reading the instructions, the subject clicks on the "Start" button and the first pair of sounds is played. A pop-up window asks "Do you hear a difference?" and the subject clicks "yes" or "no". Momentarily, the next pair of sounds

is presented, and the process repeats. The total allowable test time is 45 minutes; the interface stops running after time runs out, even if all 204 tests have not been completed. Testing occurs in three periods of 15 minutes, with a five-minute break between each. All instructions and break periods are built in, so the subject does not need to interact with an examiner during the testing (thereby eliminating one potential source of major error).

5.3. Results of Subjective Testing

The psychoacoustic test outlined in Section 5.2 was given to 12 subjects that were obtained from the Virginia Tech faculty and student body. Approximately half of the subjects work for the Vibration and Acoustics Laboratories, but few have experience performing or administering subjective tests. Four subjects were given the test twice to determine consistency between trials (for a total of 16 data sets). The number of subjects was sufficient to achieve a strong correlation between theory and the psychoacoustic test data. The test environment is depicted in Figure 5.5. The room is approximately 12-feet-by-12-feet with a carpeted floor. The walls and ceiling are covered in a burlap-type material and cover a several-inch-thick layer of fiberglass insulation. The walls and ceiling absorb most sound above 500 Hz. The computer monitor is placed on a table in the center of the room, and the user sits facing away from the door. The user has access to the headphones, a keyboard, and a trackball. The computer is positioned under and behind the table; for noise reduction, thick melamine foam covers three sides and the top of the computer. Some ambient noise enters the testing room through and around the door. Subjects were asked whether or not they noticed any bothersome sounds emanating from outside of the room. The noise, which consists primarily of occasional faint footsteps and pinging from nearby steam pipes, is not significant in magnitude. Only one subject mentioned that he could hear the background noise; he noted that it was a conversation between two people that only lasted a few seconds.

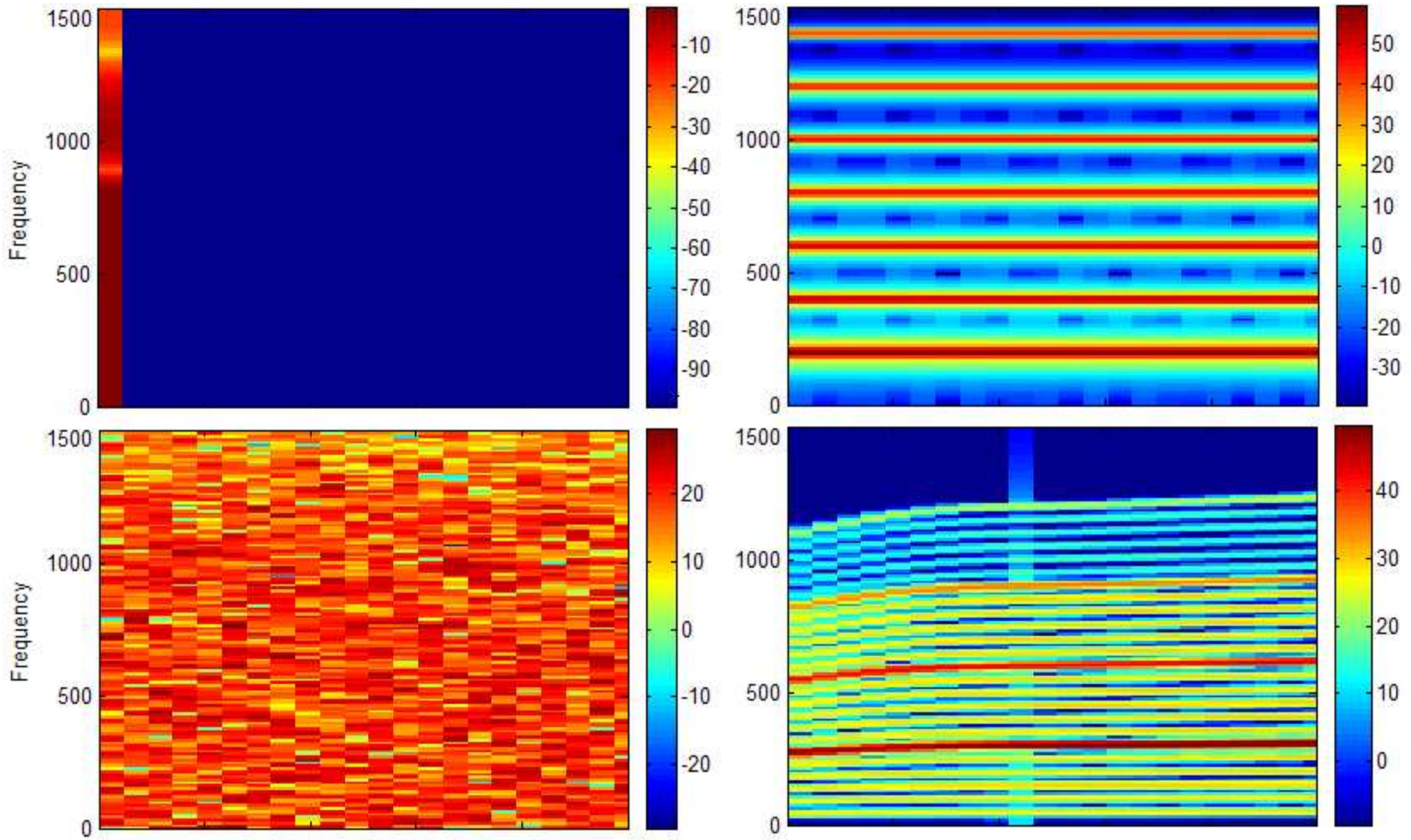

Figure 5.3. Spectograms of convolved time signals for the four noise types: Hanning pulse (top left), pure tonal (top right), white noise (bottom left), and engine noise (bottom right). Note that each plot has a different color scale. Time – from 0 to 1 second – is plotted on the horizontal axis.

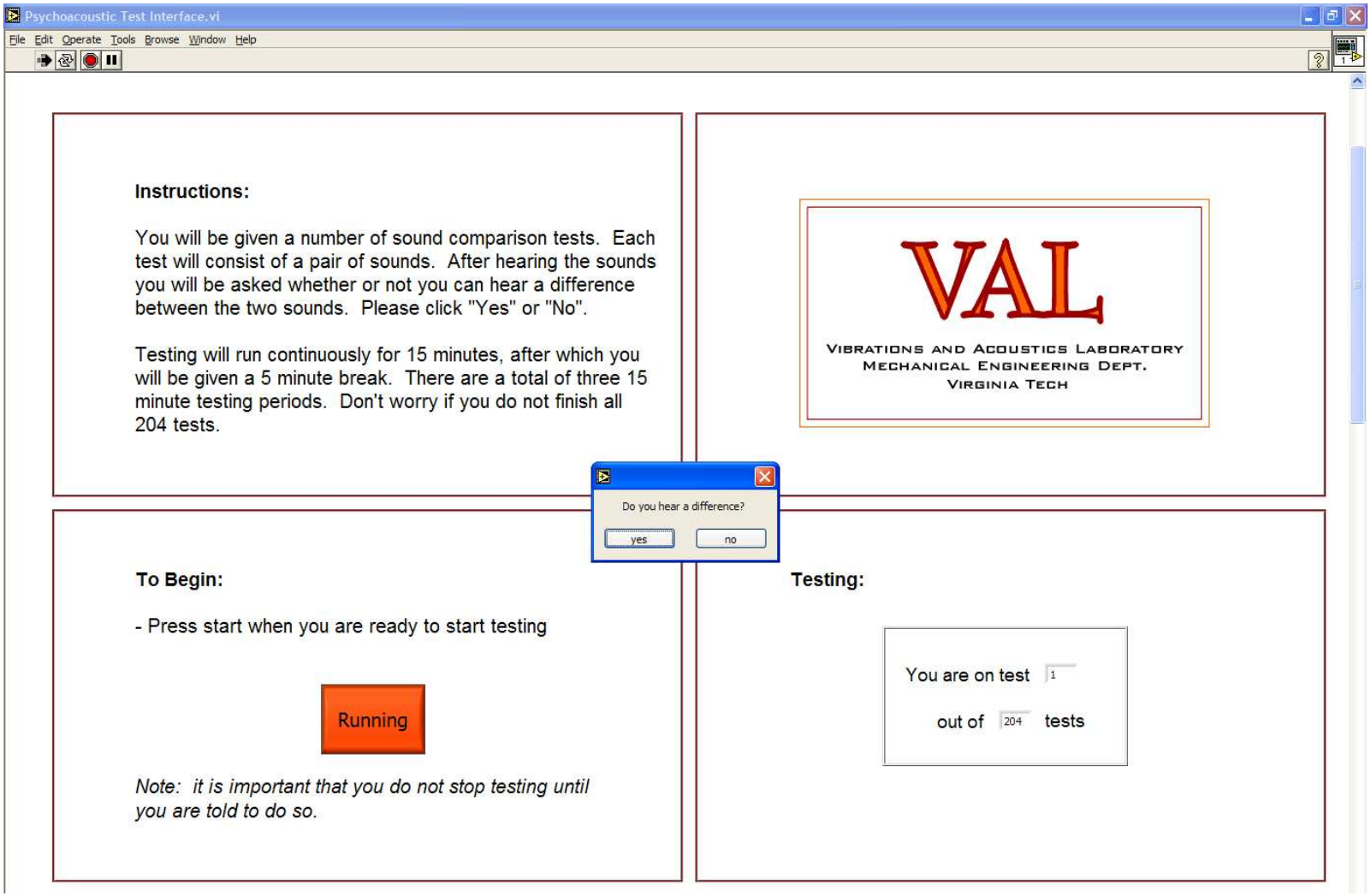

Figure 5.4. User interface for psychoacoustic testing. A test subject listens to pairs of sounds and specifies whether or not he can hear a difference between the two sounds in each pair. The instructions and break periods are built into the interface, so the subject does not need to interact with an examiner during the test.

Each subject's responses were stored in a matrix like the one shown in Figure 5.6. The test number, quality level, evaluation sphere size, noise type, and corresponding response is shown in each row. The number of sources increases from quality "1" to quality "8", and the evaluation sphere radius increases from size "1" to size "4". Noise types "1", "2", and "3" correspond to impulsive, engine, and broadband sounds, respectively. Comparisons were between the case shown in the table and the case of sphere "4" and quality "8" for the same noise type.

Figure 5.5. Subjective testing environment. The picture is taken from the door.

The 204-row matrix was then divided – based on noise type – into three sets of 68 rows. For each noise-type, three matrices were created as shown in Figure 5.7. The rows in these matrices correspond to quality levels and the columns correspond to evaluation sphere radii. The response matrix contains the sum of all of the subject's responses for each of the 32 different cases presented for the noise type. Each element corresponds to a particular case, and the value is equal to the number of times the subject indicated that he heard a difference for that case. The cases matrix contains the number of times the

subject heard each case. To obtain the percentages matrix, the responses matrix was divided (element-by-element) by the cases matrix. The percentages matrix contains the percent of times the subject could hear a difference for each case. Each subject's responses were checked for validity; this was done by making sure the number in the eighth row and fourth column did not exceed 0.33. The bottom-right corner of the matrices corresponds to a comparison of the best quality and largest evaluation sphere case versus itself. If the subject indicated that he heard a difference more than one-third of the time, his data was not included in the following analysis.

Test No.	Quality	Sphere	Noise	Response
1	1	4	3	1
2	6	2	1	0
3	5	3	1	1
4	5	2	3	0
5	6	4	2	0
6	5	3	3	1
7	5	3	2	0
8	5	3	3	1
…	**…**	**…**	**…**	**…**

Figure 5.6. Sample table of responses. A table like this one was generated as each subject performed the listening test. No difference is indicated by "0" and a difference is indicated by "1" in the response column.

Response Matrix

1	0	1	0
1	0	1	1
1	1	0	1
2	0	2	2
0	4	0	1
1	0	0	0
0	0	1	0
0	0	0	0

divided by

Cases Matrix

1	1	1	1
1	1	1	1
2	1	1	1
4	1	4	5
3	5	2	4
2	3	3	1
1	2	2	1
3	3	3	3

equals

Percentages Matrix

1	0	1	0
1	0	1	1
0.5	1	0	1
0.5	0	0.5	0.4
0	0.8	0	0.25
0.5	0	0	0
0	0	0.5	0
0	0	0	0

Figure 5.7. The subject's response data is used to generate the response, cases, and percentages matrices. The percentages matrix contains the percent of times the subject heard a difference for each case.

For each noise type, quality, and sphere size, the response matrices for all subjects were added together. The cases matrices for all subjects were also added together, and the three overall responses matrices were divided by the three overall cases matrices to

produce an overall percentage matrix for each noise type. The overall cases matrices are presented in Table 5.2. The iterative portion of the subjective testing allowed for more comparisons between sounds near the perceptual threshold region of the *NES*-r_e plane (number of equivalent sources and evaluation sphere radius plane), where it is more difficult for a listener to tell whether or not two sounds are different.

Table 5.2. Total number of trials for each noise type, quality, and sphere size. The iterative testing allowed for more trials near the perceptual threshold region of the *NES*-r_e plane.

	Impulsive				**Broadband**				**Engine Noise**			
Q1	12	12	12	12	12	12	12	12	12	12	12	12
Q2	12	12	13	15	12	12	14	12	15	16	15	15
Q3	14	18	18	23	13	13	16	13	20	24	22	28
Q4	25	35	35	33	19	21	20	20	28	37	37	39
Q5	44	42	46	39	38	38	43	52	44	36	40	37
Q6	33	28	26	28	37	38	36	33	29	22	25	23
Q7	24	20	16	17	30	29	24	22	18	18	16	14
Q8	40	37	38	37	43	41	39	40	38	39	37	36
	S1	**S2**	**S3**	**S4**	**S1**	**S2**	**S3**	**S4**	**S1**	**S2**	**S3**	**S4**

The results of the subjective testing were compared to the numerical correlations between the binaural signal pairs. The numerical correlation sounds was performed in the time domain. The correlation coefficient R between two time-domain binaural signals b_1 and b_2 is given by

$$R(b_1, b_2) = \frac{\text{cov}(b_1, b_2)}{\sigma_1 \sigma_2} \qquad [38] \tag{5.1}$$

where σ_1 and σ_2 are the standard deviations of b_1 and b_2, respectively, and cov(b_1,b_2) is the covariance, which is defined by

$$\text{cov}(b_1, b_2) = \left\langle (b_1 - \mu_1)(b_2 - \mu_2) \right\rangle \tag{5.2}$$

where μ_1 and μ_2 are the mean values of the b_1 and b_2, respectively, and $\langle\ \rangle$ is the mean operator. The numerical correlations coefficients are plotted in the top halves of Figures 5.8 (impulsive), 5.9 (broadband), and 5.10 (engine noise). The plots show the numerical

correlation coefficients between the best case (largest evaluation sphere and most equivalent sources, which corresponds to top-right corner) versus the other cases, which are indicated with a black dot. The value of *R* for the best case versus itself is always 1.0, and *R* typically decreases smoothly as the evaluation sphere shrinks and fewer equivalent sources are used. The drop-off in numerical correlation for the less accurate cases depends on the noise type. The lowest quality sound with the smallest evaluation sphere has a 90% correlation with the best case when the noise type is impulsive, but the correlation drops to 60% when the noise type is broadband. The lowest correlation for engine noise is approximately 80%. This makes sense because the numerical ESM error is greatest at higher frequencies, and the broadband noise type has the greatest high-frequency content.

The bottom halves of Figures 5.8, 5.9, and 5.10 show the perceived differences between pairs of binaural signals in the *NES*-r_e plane. The data points are denoted by black dots; the numerical values at each dot are the percentage of the trials in which the subjects indicated that they heard a difference between that case and the best case. The trends in the psychoacoustic test data and the numerical correlation plots are similar for all three noise types (the color scale is inverted between the two types of data because the top plots show correlation and the bottom plots show differences). For all noise types there are regions in the NES-re plane that correspond to binaural signals that are indistinguishable from the best case signal. This region is largest for the engine-revving sound; the value of r_e or *SPWL* can be cut by 40-50% from the best case before subjects begin to notice a difference. For the impulsive noise type, the value of r_e or SPWL can be reduced by 30-40% from the best case before subjects begin to notice a difference. Subjects are most sensitive to changes in r_e and *SPWL* when listening to the broadband signals. This makes sense because, in general, the numerical correlations are lower for the broadband signals. In the upper frequency range, LSQR error from ESM is greatest; of the three noise types, the broadband signals contain the most high frequency content.

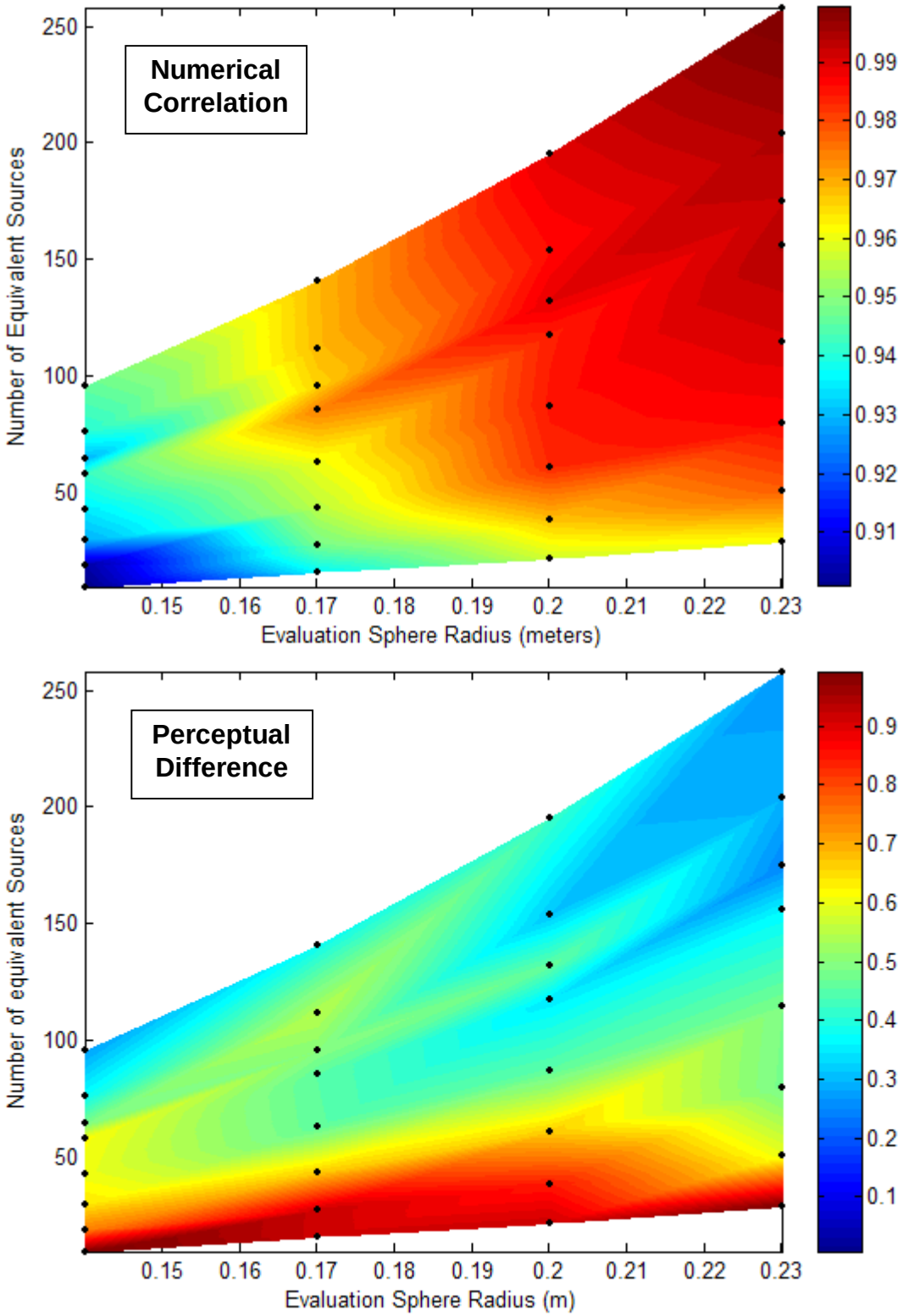

Figure 5.8. Numerical and perceptual difference between the highest quality binaural signal and sounds created with fewer equivalent sources and smaller evaluation spheres. These two plots show the results for the impulsive noise type; the black dots indicate the points that were tested.

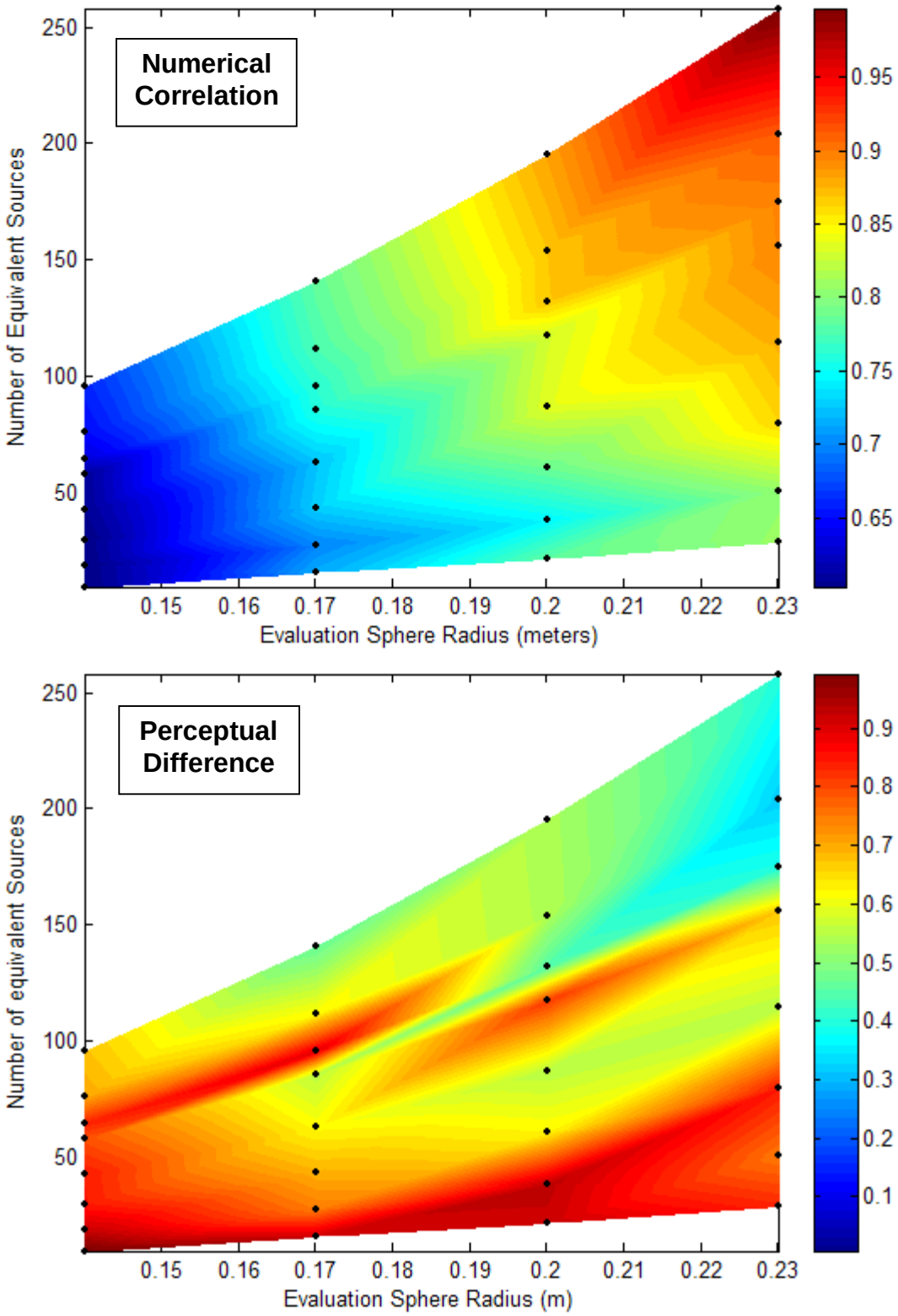

Figure 5.9. Numerical and perceptual difference between the highest quality binaural signal and sounds created with fewer equivalent sources and smaller evaluation spheres. These two plots show the results for the broadband noise type; the black dots indicate the points that were tested.

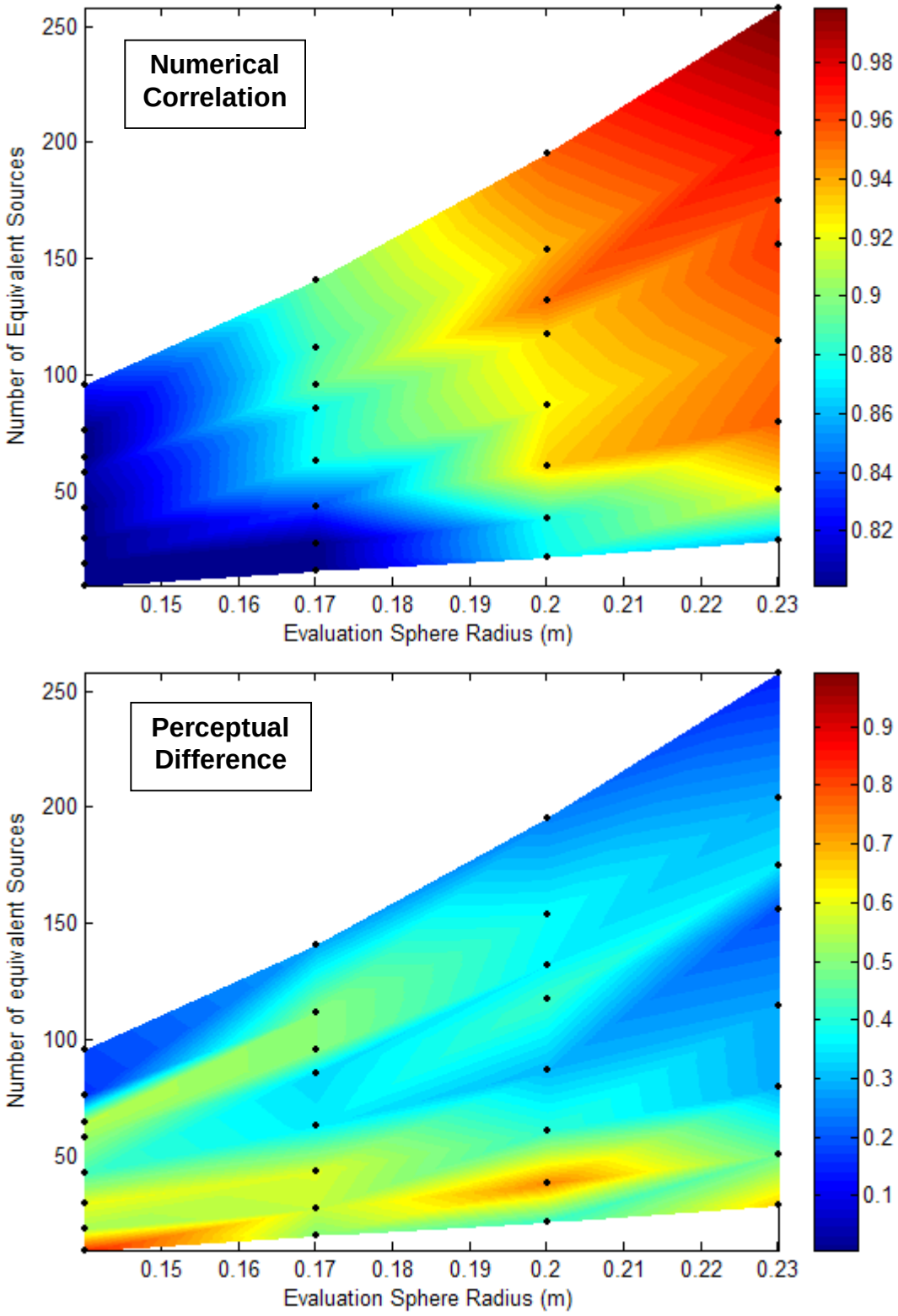

Figure 5.10. Numerical and perceptual difference between the highest quality binaural signal and sounds created with fewer equivalent sources and smaller evaluation spheres. These two plots show the results for the engine-revving noise; the black dots indicate the points that were tested.

The next step in the data interpretation is to obtain, if feasible, a linear fit between the numerical correlations *N* and the perceived differences *D* for each noise type. This curve will allow a structural designer or psychoacoustic test proctor to predict what percentage of subjects will hear a difference between two signals with a known numerical correlation and noise type. The 32 data points – indicated by the black dots in Figures 5.8, 5.9, and 5.10 – were used to determine the coefficients a_0 and a_1 in the linear regression model

$$D = a_0 + a_1 N \tag{5.3}$$

The linear fits for each noise type are plotted in Figure 5.11. Thirty-two blue dots mark the "perceived difference versus numerical correlation" data points in the *N*-*D* plane; the magenta lines were calculated using by applying a least-squares regression algorithm. A straight line appears to be the proper curve for the data. Next, the correlation coefficients and P-values were calculated for the *N*-*D* pairs of each noise type. A P-value is the probability of obtaining a correlation coefficient as high as the calculated correlation coefficient by random chance (when the true correlation is actually zero) [39]. As a general rule, two data sets are significantly correlated if the P-value is less than 0.05; this is true for all three noise types. The correlations of 59% (impulsive), 68% (broadband), and 60% (engine noise) are statistically significant. If the noise type is known, then the model introduced in Equation 5.3 can be used to predict a perceptual difference between two sounds in the presence of a numerical correlation. All linear regression and correlation coefficients, as well as P-values, are summarized in Table 5.3.

Table 5.3. Linear fit and correlation coefficients for the comparison of the perceptual differences to the numerical correlations.

	a_0	a_1	R	P-value
Impulsive	4.94	-4.55	0.59	0.000419
Broadband	1.49	-1.04	0.68	0.000022
Engine Noise	1.49	-1.23	0.60	0.000282

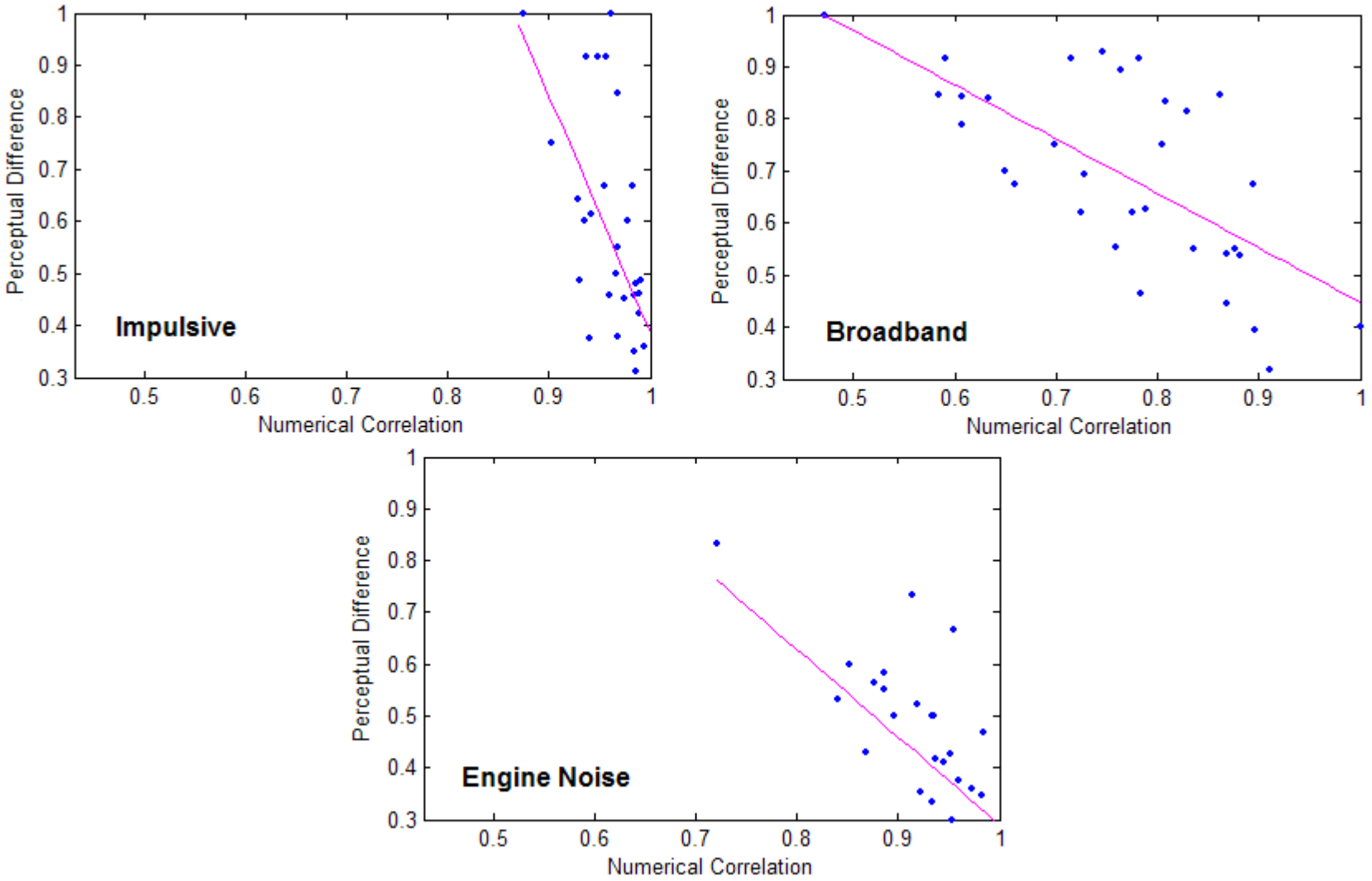

Figure 5.11. Linear fit between the perceptual difference and numerical correlation in the comparison of binaural signal pairs. A strong link exists between a numerical difference and a perceivable difference between two sounds of the same type.

In addition to the impulsive, broadband, and engine-revving noises, a purely-tonal noise type was psychoacoustically tested using four subjects; the spectrogram for this noise is plotted in the top-right corner of Figure 5.3. The pure tone contains a 200-Hz component and smaller amounts of the next six harmonics. Thirty-two binaural signals were created using the same values of r_e and *NES* as for the other three noise types, and the same comparisons were presented to the subjects. The purely tonal noises were mixed with impulsive and broadband signals to prevent the subject from learning the test. Figure 5.12 shows the perceptual difference between the best cases and the other cases (indicated by black dots); the numerical correlation plot is nearly identical to the plot for the engine-revving noise (top of Figure 5.10). However, subjects were able to consistently tell a difference between all pairs of binaural signals except for the trials of the best cases versus itself. The numerical ESM error is not large when it is averages over the entire 1000 Hz octave band, but the error at each frequency can be large; only seven frequencies are present in the signals. The perceived differences in the binaural signals are large for all comparisons except for the best case versus itself, which means that humans are sensitive to moderate changes in the proportions of the various tones that are present in a sound. Even the best case may not be accurate; if the evaluation sphere size was increased and more equivalent sources were added, the binaural signal would likely sound different from the current best case.

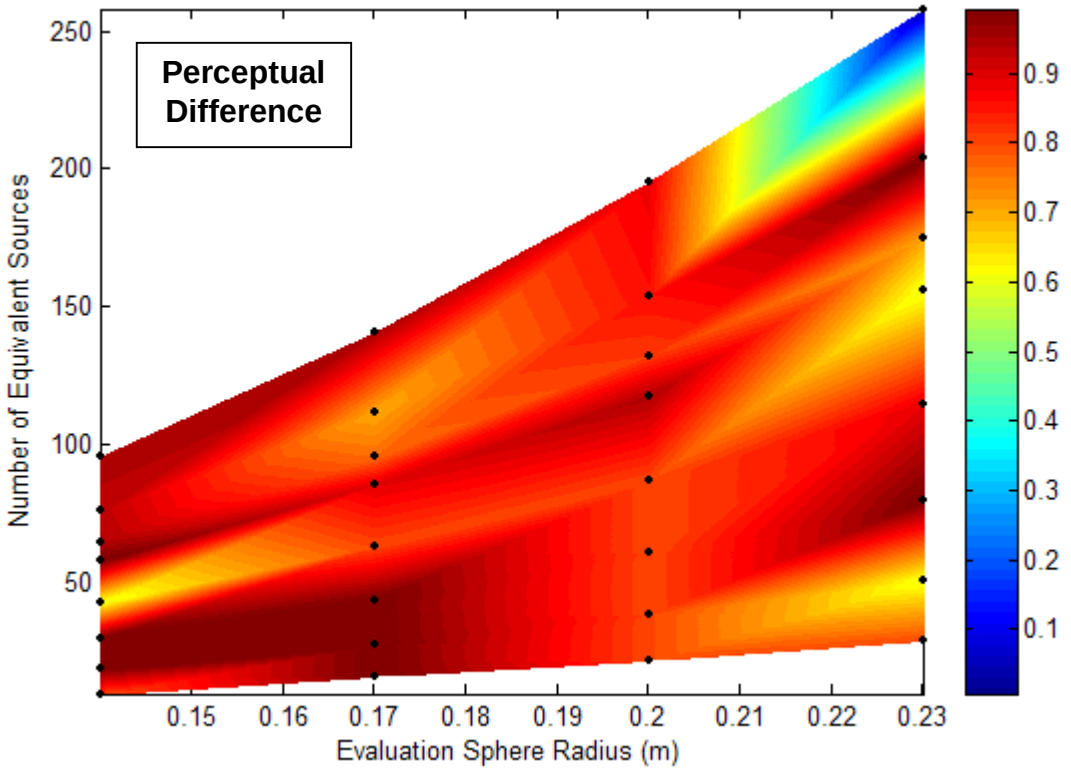

Figure 5.12. Numerical and perceptual difference between the highest quality binaural signal and sounds created with fewer equivalent sources and smaller evaluation spheres. These two plots show the results for a purely tonal noise; the black dots indicate the points that were tested.

The plots in Figures 5.8, 5.9, 5.10, and 5.12 indicate that the robustness of ESM, when applied to the finite element model output of a simple structural design, is heavily dependent on what type of noise signal is exciting the structure. For impulsive or broadband excitation of the structure, using only 3.7 equivalent sources per acoustic wavelength is adequate from a perceptual standpoint. When the input has frequency content that is changing in time, as is the case in the engine-revving noise, then a binaural signal using only 2.5 sources per wavelength may be indistinguishable from a signal that uses 4.5 sources per wavelength. For all three types of excitation, a point of diminishing return on the evaluation sphere radius exists. Larger evaluation spheres require more equivalent sources to recreate their internal sound field, which translates to a greater computational power requirement; there is typically an upper limit on the number of sources that can be handled by the auralization hardware. Excitation that contains only a

few frequencies that are not changing in time produce a case where ESM may not be feasible. No point of diminishing return on the evaluation sphere size or the number of equivalent sources was determined for the purely tonal noise type; at least 4.5 sources per wavelength or a huge evaluation sphere must be used.

For three noise types, the relationship between the numerical correlation of two binaural signals and the percentage of subjects that will perceive a difference between the sounds was established. If the noise type and numerical correlation between two sounds is known, then the ability of subjects to perceive a difference can be predicted. If ESM is used in the psychoacoustic evaluation of a simple structural design, then an appropriate number of equivalent sources can be chosen in order to produce the desired numerical and perceptual accuracies.

Chapter 6: Conclusion

This thesis summarized the author's work in the field of virtual acoustics. The equivalent source method – which is integral to the real-time, head-tracked auralization of a modeled structure – was evaluated both numerically and psychoacoustically. Section 6.1 provides a summary of the first five chapter of this thesis and reiterates the important conclusions that were drawn. Section 6.2 suggests areas for future work that related to virtual acoustic prototyping.

6.1 Summary and Conclusions

This thesis began with an introduction to virtual acoustics and a description of the research that was performed. Chapter 1 started with a definition of virtual acoustic prototyping; the definition was followed by a description of head-related transfer functions (HRTFs), the equivalent source method (ESM), and the use of finite element modeling (FEM) in the auralization process. The motivation for virtual acoustic prototyping was also established; it can be a useful tool in the design and subjective evaluation of a structure's acoustics. Engineering firms can save a lot of money by auralizing and enhancing the acoustics of an automobile or office building during the design phase (rather than after a physical prototype is built). Chapter 1 also contained a review of recent literature on virtual acoustics, HRTFs, ESM, the use of FEM in structural acoustic problems, and psychoacoustics. A tradeoff exists between numerical accuracy and computational time. The scope of the work in this thesis was defined; the goal was to subjectively validate ESM and the subsequent application of HRTFs by determining what numerical error level is acceptable.

Chapter 2 introduced the basic theory behind the virtual acoustic prototyping method used in this thesis. HRTFs and ESM were explained in detail; immediately following was a discussion of the numerical error associated with an auralization. Error arises from three sources – the evaluation volume size, the number of equivalent sources, and the interpolation and extrapolation of HRTFs. Using a smaller evaluation sphere means that a smaller portion of the sound field from a vibrating structure is recreated with

equivalent sources; the smaller sound field may not be representative of the full field in and around the structure. Using too few equivalent sources prevents the sound field inside the evaluation sphere from being recreated accurately. The final step in the auralization process is the creation of a binaural signal from the set of equivalent sources. HRTFs are only measured at discrete azimuth and elevation angles, and they are usually not measured for elevation angles below -40°. If an equivalent source does not lie at the exact azimuth and elevation of a measured HRTF, the appropriate HRTF must be estimated from the known HRTFs. Each person has a different set of HRTFs, so using a representative set during auralization introduces some error. The goal of this thesis is to psychoacoustically determine what numerical error level is noticeable.

The numerical modeling of the virtual acoustic prototyping process was covered in Chapter 3. The finite element method was introduced; the theory behind the method was discussed, and to demonstrate the method, the matrices were formulated for a rod in compression. The structural acoustic test platform used in this thesis was introduced. FEM was performed on the enclosure; the acoustic space, Lexan window, and MDF walls were all modeled in an effort to produce an satisfactory match between the acoustic responses of the modeled structure and a physical prototype of the same structure. ESM was applied to the model output to recreate a portion of the sound field in the enclosure; four different evaluation sphere radii and eight values of *SPWL* (equivalent sources per acoustic wavelength) were used, which provided a total of 32 cases. HRTFs were applied to each equivalent source set, and an analysis of the numerical error was performed. A large evaluation sphere should be used whenever possible, but there is a diminishing return on the use of more equivalent sources; above approximately *SPWL* = 3.0, more equivalent sources does not significantly reduce the numerical error. This observation agrees with theory.

The structural acoustic test platform modeled in Chapter 3 was constructed in Chapter 4. The assembly of the panels was discussed in detail. Microphones and accelerometers were added to the enclosure and walls, and a shaker with force transducer was attached to the Lexan window to excite the box. Frequency response functions (microphone output per shaker input) were measured at 472 points to determine the dominant acoustic natural frequencies between 50 and 250 Hz. The material properties in

the model were adjusted within reason to create a match between the first six acoustic modes. The model was solved again between 50 and 1500 Hz, and the MDF damping coefficients were changed slightly to provide a satisfactory match between the numerical model and the physical prototype. A 36% correlation was achieved; this is numerically low, but the sounds of the modeled acoustic field and the measured acoustic field are very similar. Finally, the Knowles Electronic Manikin for Acoustic Research (KEMAR) was inserted into the enclosure; the response at its left and right ears was measured for a shaker force input. The measured responses were compared to the binaural signals calculated in Chapter 3; once again, the numerical difference was large, but the perceived difference between the sounds was much smaller. The numerically-large discrepancies occur as a result of small deviations in the material properties and boundary conditions from their modeled values. Few of the high-frequency modes overlap; but the two very-different frequency responses can occur in two very similar instances of the same structure. This result is consistent with the available literature.

Chapter 5 covered the psychoacoustic evaluation of ESM and the application of HRTFs. A comparison test was created in which subjects were asked to compare 204 pairs of sounds. The sounds were created from the equivalent sources (which are impulse responses). The impulse responses were convolved with four types of sounds – a Hanning pulse, white noise, engine noise, and a series of harmonics. Subjects were asked whether or not they could hear a difference between the best binaural signal for a given noise type (created using the largest evaluation sphere and the most equivalent sources) and a lower quality sound of the same noise type (created using a smaller evaluation sphere and/or fewer equivalent sources). The goal was to determine how many equivalent sources should be used to provide acceptable perceptual accuracy. A relationship between evaluation sphere size, number of equivalent sources per acoustic wavelength, and numerical correlation and perceptual differences between two binaural signals was established. Finally, a linear regression was performed to allow a structural engineer or a psychoacoustic examiner to predict the percentage of subjects that will hear a difference between two sounds if their noise type and numerical correlations are known.

For three noise types (impulsive, broadband, and engine noise), the relationship between the numerical correlation of two binaural signals and the percentage of subjects

that will perceive a difference between the sounds was established with correlations coefficients of 59%, 68%, and 60%, respectively. If the noise type and numerical correlation between two sounds is known, then the ability of subjects to perceive a difference can be predicted with accuracy. If ESM is used in the psychoacoustic evaluation of a simple structural design, then an appropriate number of equivalent sources can be chosen in order to produce the desired numerical and perceptual accuracies; the number of sources depends heavily on the noise type. Subjects are most sensitive to changes in the number of sources when the noise type is broadband; at least 3.5 *SPWL* should be used. For impulsive and engine-revving sounds, as few as 3.0 and 2.5 *SPWL* can be used, respectively. The results of this study indicate that only a small percentage of subjects will be able distinguish these less accurate sounds from those sounds which use 4.5 *SPWL*.

6.2 Future Work

The finite element method has a very limited capability to predict the response of a coupled structural-acoustic model. To date, ESM has only been applied to the output of a finite element model. The response of a structure must be calculated for a discrete set of frequencies in a range of interest, and the characteristic length of each mesh element must be less than one-sixth of an acoustic wavelength. Computation times are on the order of days or weeks for models as large as the one used in this thesis. During the design and prototyping phase of an automobile or building, long periods of time are not always available. Therefore, as a matter of future work, ESM should be applied to other structural acoustics modeling techniques such as ray tracing, statistical energy analysis (SEA), and the recently-developed hybrid FEM-SEA method.

In this thesis, ESM was only validated for short samples of three noise types. The subjective testing and analysis described in Section 5.3 should be extended to other types of noises during future research endeavors. Not much is known about why certain noise types can be rendered accurately using fewer equivalent sources (and why ESM appears not to work for purely tonal noises) when all of the binaural signals were created from the same finite element model output. A set of general guidelines should be established to

determine – without psychoacoustic testing – how many equivalent sources should be used to render new noise types.

The subject testing environment used in this thesis did not accurately represent the surroundings of the structure being tested. A person's ability to distinguish between two binaural signals and a sound's impression on a listener are greatly affected by the testing environment. To date, some work has been done at the NASA-Langley Research Center to create a aircraft simulator for use in subjective evaluation of sound [27]. An extensive literature search revealed that a realistic automobile simulator has not been created for use as a psychoacoustic test platform. This should be performed as a matter of future work; a subject's ability to evaluate the acoustics of a car or truck would be enhanced in a realistic driving environment.

None of the psychoacoustic testing in this thesis was performed in a real-time, head-tracked format. If the auralization software is working correctly, a subject should not be able to tell the difference between a binaural signal that is rendered in real time and a signal that has been preprocessed. However, if a head tracker is used, the subject can rotate his head and the sound source will still appear to be coming from the same direction in space; the binaural signal changes as the head rotates and translates. A head tracker can only be used properly in real-time. Head tracking introduces greater complexity to the auralization process. ESM should be validated using a real-time, head-tracked subjective testing in future experiments to make sure the method does not lose its robustness with the additional complexities.

This thesis showed that ESM is a useful tool in the design and psychoacoustic evaluation of coupled structural-acoustic systems. As a matter of future research, the work outlined here should be performed to validate ESM under a wider variety of circumstances. Following further validation, an effective auralization tool can be placed in a user-friendly package and distributed to architects and structural engineers; this project is currently being undertaken by AuSIM, Inc. in Mountain View, California.

References

1. Kleiner, M., B.I. Dalenback, and U.P. Svensson, "Auralization—an overview," *Audio Engineering Society* (1993). (Presented at the *91st AES Convention*, October 4-8, 1991, New York; revised October 13, 1993.)
2. Sandberg, G. and P. Goransson, "A Symmetric Finite Element Formulation For Acoustic Fluid-Structure Interaction Analysis", *J. Sound Vib.* **123**, (1988), p. 507-515.
3. Lalime, A.L., "Development of a Computationally Efficient Binaural Simulation for the Analysis of Structural Acoustic Data", *Digital Library and Archives: Virginia Polytechnic Institute and State University*, (2002) < http://scholar.lib.vt.edu/theses/available/etd-08142002-122633/ >.
4. Cheng, C.I. and G.H. Wakefield, "Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFS in Time, Frequency and Space", *J. Audio Eng. Soc.* **49**(4), (2001), pp. 231-246.
5. Bedard, P., "Sustainable Sports Sedans", *Car and Driver* (October 2004), <http://www.caranddriver.com/article.asp?section_id=15&article_id=8550>.
6. Quiroga, T., "21st Century Muscle Cars", *Car and Driver* (January 2005), <http://www.caranddriver.com/article.asp?section_id=15&article_id=8908>.
7. Wenzel, E.M., "A System for Three-Dimensional Acoustic 'Visualization' in a Virtual Environment Workstation", *IEEE* (1990).
8. Johnson, M., A. Lalime, F.W. Grosveld, S.A. Rizzi and B.M. Sullivan, "Development of an efficient binaural simulation for the analysis of structural acoustic data," *Proc. 8th Conf. Rec. Adv. Struct. Dyn*, Paper #77, (July 2003)
9. Koopmann, G.H., L. Song, and J.B. Fahnline, "A method for computing acoustic fields based on the principle of wave superposition", *J. Acoust. Soc. Am.* **86**(6), (1989), pp. 2433-2437.
10. Bianchi, F.W. and R.H. Campbell, "The Virtual Orchestra: Technical and Creative Issues", *J. Sound Vib.* **232**(1), (2000), pp. 275-279.

11. Gladwell, G.M.L. and G. Zimmermann, “On Energy and Complementary Energy Formulations of Acoustic and Structural Vibration Problems”, *J. Sound Vib.* **3**(3), (1966), pp. 233-241.
12. Gladwell, G.M.L., “A Variational Formulation of Damped Acousto-Structural Vibration Problems”, *J. Sound Vib.* **4**(2), (1966), pp. 172-186.
13. Petyt, M., J. Lea, and G.H. Koopmann, “A Finite Element Method For Determining the Acoustic Modes of Irregular Shaped Cavities”, *J. Sound Vib.* **45(**4), (1976), pp. 495-502.
14. Craggs, A. and G. Stead, “Sound Transmission Between Enclosures – A Study Using Plate and Acoustic Finite Elements”, *Acustica* **35**(2), (1976), pp. 89-98.
15. Craggs, A., “The Application of Acoustic and Absorption Finite Elements to Sound Fields in Small Enclosures”, *Finite Element Applications in Acoustics* (1981).
16. Doyle, G.R., and Faulkner, L.L., “Three-Dimensional Finite Element Acoustic Pressure Analysis with Complex Impedance”, *Finite Element Applications in Acoustics* (1981).
17. Nefske, D.J., A. Wolf, Jr., and L.J. Howell, “Structural-Acoustic Finite Element Analysis of the Automobile Passenger Compartment: A Review of Current Practice”, *J. Sound Vib.* **80**(2), (1982), 247-266.
18. Mace, B.R. and P.J. Shorter, “Energy Flow Models from Finite Element Analysis”, *J. Sound Vib.* **233**(3), (2000), pp. 369-389.
19. Hopkins, C., “Vibrating transmission between coupled plates using finite element methods and statistical energy analysis”, *Applied Acoustics* **64**, (2003), pp. 955-973.
20. Shorter, P., V. Cotoni, and R. Langley, “Numerical and Experimental Validation of Hybrid FE-SEA Method”, *NOISE-CON 2004*, Baltimore MD, (July 12-14, 2004).
21. Mahn, J.P. and P.V. Bayly, “Impact Testing of Stereolithographic Models to Predict Natural Frequencies”, *J. Sound Vib.* **224**(3), (1999), pp. 411-430.

22. Song, L., G.H. Koopmann, and J.B. Fahnline, "Numerical errors associated with the method of superposition for computing acoustic fields", *J. Acoust. Soc. Am.* **89**(6), (1991), pp. 2625-2633.
23. Jeans, R. and I.C. Mathews, "The wave superposition method as a robust technique for computing acoustic fields", *J. Acoust. Soc. Am.* **92**(2), (1992), pp. 1156-1166.
24. Esteve, S. and M. Johnson, "Fourth Contract Report", *Improving Binaural Simulation of Structural Acoustics*, (November 2004), submitted to AuSIM, Inc. in Mountain View, CA.
25. Glorig, A., *Audiometry: Principles and Practices*, Baltimore: The Williams & Wilkins Company (1965).
26. Zwicker, E. and H. Fastl, *Psychoacoustics*, Berlin: Springer-Verlag (1999).
27. Sullivan, B.M. and C.A. Powell, "A new aircraft interior noise simulator for psychoacoustic testing", *The Journal of the Acoustical Society of America* **102**(5), (1997), p. 3191.
28. Rayleigh, L., "On Our Perception of Sound Direction", *Philosoph. Mag.* **13**, (1907).
29. Russell, D., "Sound Fields Radiated by Simple Sources", *Acoustics Animations*, (2001), < http://www.kettering.edu/~drussell/Demos/rad2/mdq.html >.
30. Paige, C.C. and M.A. Saunders, "LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares", *ACM Transactions on Mathematical Software* **8**(1), (1982), pp. 43-71.
31. Esteve, S. and M. Johnson, "Third Contract Report", *Improving Binaural Simulation of Structural Acoustics*, (September 2004), submitted to AuSIM, Inc. in Mountain View, CA.
32. Mennitt, D. and M. Johnson, "Determination of the minimum size of the evaluation sphere in order to achieve accurate local simlation", *Not yet published*, Blacksburg, Virginia: Vibration and Acoustics Laboratories, Virginia Tech, (2005).

33. Gardner, B. and K. Martin, “HRTF Measurements of a KEMAR Dummy-Head Microphone”, *MIT Media Lab Machine Listening Group*, (2000), < http://sound.media.mit.edu/KEMAR.html >.
34. Roginska, A., Mountain View, California: AuSIM, Inc., *E-mail consultation,* (2005).
35. Bickford, W.B., *A First Course in the Finite Element Method*, ISBN 0-256-07973-0, pp. 4-29.
36. The MathWorks, “Rational Fraction Polynomial Method” , *Matlab Central – An open exchange for the MATLAB and Simulink user community,* < http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=3805&objectType=file >.
37. Inman, D.J., *Engineering Vibration*, Upper Saddle River: Prentice Hall (2001), pp. 293-294, 533-563.
38. Weisstein, E.W., "Covariance", *From MathWorld--A Wolfram Web Resource,* < http://mathworld.wolfram.com/Covariance.html >.
39. The MathWorks, “corrcoef (MATLAB Functions)”, *The MathWorks – Accelerating the pace of engineering and science*, < http://www.mathworks.com/access/helpdesk/help/techdoc/ref/corrcoef.html >.

Appendix A: Microphone Array Measurements

The planar microphone array shown in Figure 4.7 was primarily used to determine the acoustic modes of the structure between 50 and 250 Hz. The array was configured to allow 59 microphone measurements (on a two-dimensional grid) to be taken rapidly. To A three-dimensional grid was achieved by sliding the array to eight different depths in the enclosure. A total of 472 frequency response functions (FRFs) were measured for two forcing functions:

- a shaker force input to the window
- a speaker excitation in the corner of the enclosure

The speaker FRFs were used to determine the most dominant natural frequencies and the corresponding mode shapes. In ABAQUS, the material properties of the acoustic space were tweaked to provide a better match. Figure A-1 shows a sample comparison of the acoustic mode shapes calculated in ABAQUS and the mode shapes determined from the measured data.

After the low-frequency acoustic modes were matched, then the damping parameters in the MDF and Lexan panels were modified, within reason, to give a better match at higher frequencies. Obtaining the overall match between the model and enclosure was an iterative process, and relied on the data from the shaker excitation of the structure. The data collected with the microphone array were compared to the final finite element model to determine a 36% correlation.

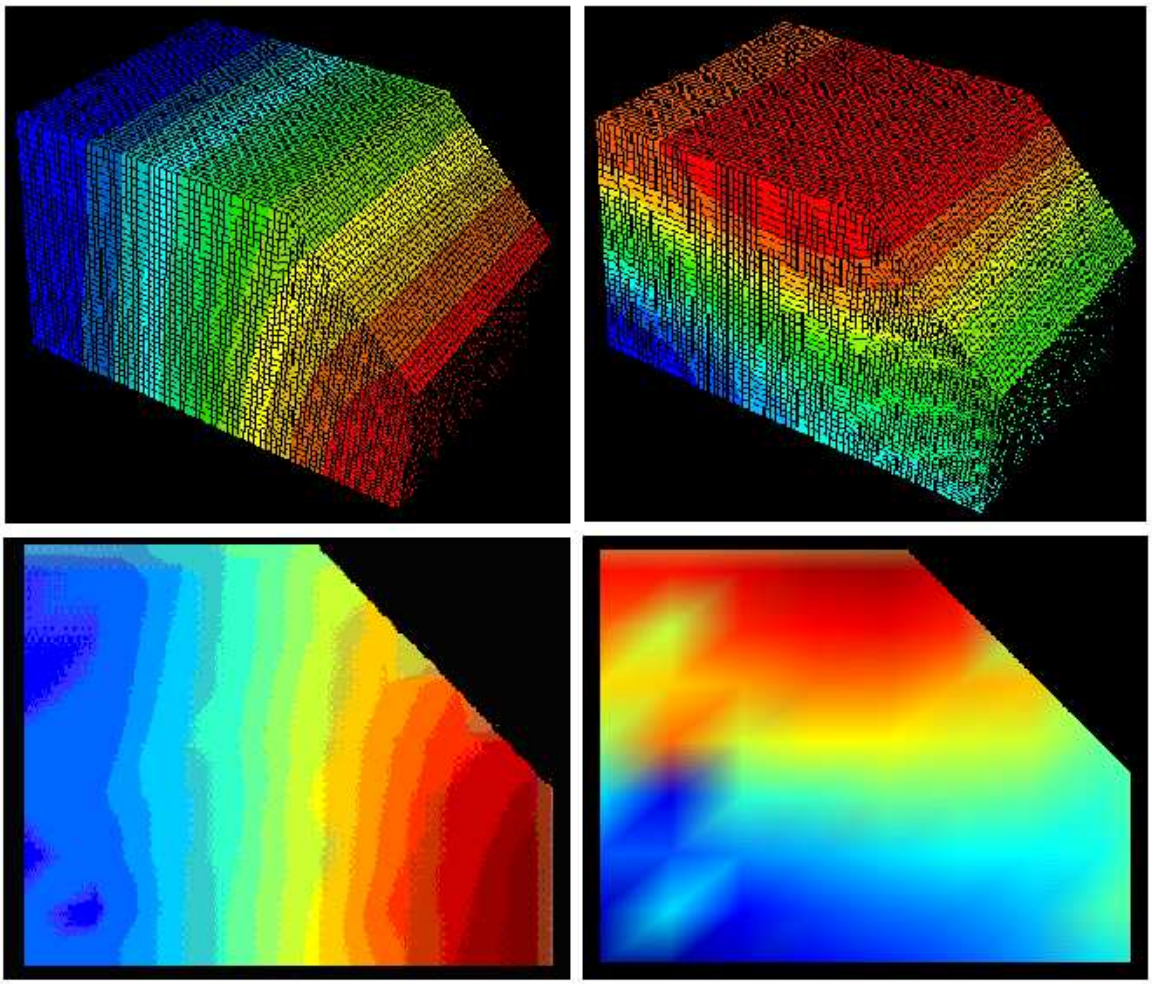

Figure A-1. Comparisons for two low-frequency acoustic modes that appear in the finite element model and microphone array data. The top plots are from ABAQUS and the bottom plots were generated from the measured FRFs in Matlab. The left plots are for the mode at 102 Hz an the right plots are for the mode at 155 Hz.

Vita

Zachary T. Carwile was born in Richmond, VA on September 29, 1982 to Linwood and Brenda Carwile. He graduated *Magna Cum Laude* with B.S. in Mechanical Engineering from Virginia Tech in May 2004. In August 2004, he began the pursuit of his M.S. in Mechanical Engineering at Virginia Tech, with a concentration in acoustics and signal processing. Following graduation, he will begin working in the Engineering Development Group at The MathWorks in Natick, Massachusetts. In addition to most things science and engineering-related, his interests include outdoor athletic activities such as ultimate frisbee and mountain biking.