THE USE OF SOUND INTENSITY FOR CHARACTERISATION OF REFLECTED ENERGY IN SMALL ROOMS

Julián ROMERO-PÉREZ

School of Computing, Science and Engineering University of Salford, Salford, U.K.

Submitted in Partial Fulfilment of the Requirements of the Degree of Doctor of Philosophy, May 2013

Table of contents:

List of figures

List of tables

Personal motivation

The first motivation for this research was born because of some discrepancies found when doing some reverberation time (*RT*) measurements. At the time, I was comparing data measured against *RT* calculations done with ray-tracing acoustic software called *Odeon.* It was clear that the measurement results presented in my M.Sc. dissertation back in 2007 were inaccurate; consequently, I wanted to understand how to attain good acoustic measurements. First question asked myself was: how can I contribute to the study of small spaces, which do not have a diffuse field? I remembered I was intrigued with the question of existence of a measure for diffuseness and how to use it on room acoustics. The literature review found on that topic was scarce. At the end of my M.Sc. Audio Acoustics at the University of Salford, I found myself dissatisfied with the level of knowledge that I grasped during my courses. I needed to continue further because I was pursuing deeper understanding of acoustics. I remember reading the Ph.D. thesis of Dr Bruno Fazenda who later became my main supervisor. Reviewing all the subjects studied in the M.Sc., I was totally into room acoustics. Amongst all the Ph.D. theses that I read during my M.Sc. in Salford, the one authored by Dr. Fazenda was my favourite because it combined the critical listening spaces and low-frequency characterisation. Coincidentally, when I was finishing my dissertation, he was in search of supervision of a Ph.D. student. Better yet, the topic of research was very interesting and challenging because it was mixing small rooms and the other part was not touched in his thesis, i.e. the effect of mid and high frequencies. We had a mutual supervisor named Dr. William J. Davies, who recommended me to Dr. Fazenda to obtain that position through a fee waiver at the University of Huddersfield. The idea to combine the study of small room acoustics and the characterisation of reflected energy attracted my attention because nowadays, the majority of the designs that can be done practically are small spaces. The common practice to characterise small rooms using a single omni-directional microphone lacks any directional information. Therefore, this topic was a perfect continuation of the studies of non-diffuse spaces, which was the subject of my M.Sc. Audio Acoustics dissertation. It

was a great area to explore because of the scarce information found on the topic and my ultimate goal, which was to focus on critical listening spaces.

Acknowledgments

I would like to thank my supervisor, Dr. Bruno M. Fazenda, for his invaluable support throughout the development of the project, and his wise comments. He recognised my determination for studying acoustics, and helped me to accomplish my studies. I thank him so much for helping me to be transferred from the University of Huddersfield to the University of Salford to complete the last part of my Ph.D. I also want to express gratitude to my second supervisor, Dr. Mark Avis, for his unconditional and continuous support, motivation and deep insight on transducers and microphone arrays to detect sound reflections. During my former years at the University of Huddersfield, I received a great deal of feedback from Dr. Luyang Guan and Dr. Hidajat Atmoko, both of whom paid me a lot of attention and gave me fruitful explanations on the digital signal processing (*DSP*) analysis and the development of a custom microphone array. More importantly, they showed me what research is all about, and what the practical extents on a project are. I also recognise the valuable vision of Professor Jamie Andrea Shyla Angus for providing me the support and criticism during my first and second *IOA* conferences and for his worthy comments on the issues of B-Format limitations on reflection detection, which was influential for focusing on p-p intensity techniques. I truly thank Dr. David Griesigner for his early recommendation to use sound intensity for the detection of sound reflections as a comment after presenting my first paper of *IOA 2008*. Likewise, Dr. Damian Murphy was very helpful with guidance on short Fourier transform applied to measuring early reflections.

I dedicate this thesis to my daughter Victoria Romero Fernández and Dr. Eulàlia Peris Aguiló. I admire them both as unstoppable, beautiful human beings who have always inspired me to fight for what I truly want, and to always offer the best, no matter how difficult the circumstances are. I wish to thank my ex-wife Diana Fernández Treviño for all her temporary love and unconditional support during the early years in U.K., when we

I want to thank to José Antonio Almagro Pastor for his interesting comments on comparison of p-p intensity probes and p-u probes. I want to acknowledge Dr. Javier Redondo for his deeply discussions regarding 3-D scattering and 3-D diffusors.

I want to thank Dr. Hyok Jeong for his help on wave theory and reflection modelling, his human kindness to recognize my early work and his prays. I truly thank Dr. Hyun Lim for his patience and countless discussions about intensity and particle velocity nature at late night at the office.

I thank Dr. Jonathan Wakefield for supporting me while I was giving some laboratory classes at University of Huddersfield, to Dr. David Moore for their guidance on B-Format and his introduction to *DirAC*, to Steven Fenton and Mark Mynett for his comments on studio monitors, recording and production and digital signal processing while we were studying in the University of Huddersfield.

I want to thank my former M.Sc. colleague Robert Will for very interesting discussions on modal decomposition for enclosures and transducers theory, which has helped my understanding of small room acoustics.

Irrevocably, I want to express my deepest thanks to all the staff that was in the Post-Graduate Office of the Acoustics Research Department at the University of Salford. Being there, it truly has held very stimulating acoustic discussions at any minor provocation. They helped me to balance my academic life with a weekly social gathering every Friday at the Crescent. People that I can recall that enjoyed together are: Dr. Andrew Elliot, Dr. Fouad Bechwati, Dr. Konstantinos Dadiotis, Dr. Jonathan Hargreaves, Dr. Paul Kendrick, Dr. Rodolfo Venegas Castillo and his wife Yasna Aguilera, Dr. Richard Hughes, Dr. Charlie Mydlraz, Dr. Neil Bruce, Dr. Benjamin Piper, Dr. Tomos Pentre, Dr. Jenna Condie, Dr. Elisabeth Duarte, Dr. Anton Krykin, Dr. Liam Kelly, Dr. Diego Turó, Claire Churchill, Tony Letakin, Linda McKittrick Gedemer, Natalia Szczepancyk, Masimo Serafini, Gennaro Sica, Zbigniew Koziel, Steven Wheeler, Dr. Robert Oldfield, Calum Sharp, Andrew King, Tobias Alexander Ackroyd, Michael Hudelmaier, and Rachel Bennett.

Lastly but not at last, I thank all the personnel from the University of Salford especially to Claire Lomax, Danny McCaul, who have been very kind. They trusted me, provided me expensive equipment, and helped me with technical issues. I thank them for let me booking the semi-anechoic chamber for countless experiments.

To continue thanking more people at the end of this section, I want to thank to all my friends not related to the academic team of Acoustics, which are amazing musicians and allowed me to play wonderful funk music with the band, "The Scratters and the funky disco pig". That creative activity helped to mix the academic side with the music scene on both brain hemispheres and relax from the intense work involved in my Ph.D. These artistic people are: Martyn Cawthorne, Minos Romanos, Manos Sarantidis, Biff Roxby, Adrian Drinkwater, Martin Coles, Adam Khomsi, Bob Marsh, Francis Lydiatt, Colin Crichton and Carlos Martínez.

Finally, I have to thank all my friends back in Mexico, which have contributed in one way or another to my completion of this thesis: Akinori Tago Mendoza who introduced me to audio and recording studios and music bands, Hector Castañón who introduced me to Acoustics, Azucena Marín Zamora, who has been supporting me throughout the completion of this manuscript, and Mónica Murguía Sicard for her support in my return to Mexico to face my divorce and encourage to become a Ph.D. An without doubt, I need to stress a special thanks to Alfonso Ruiz Palacios Quintero, my brother in law who brought this pieces of work to U.K. while I was troubled in legal issues with my divorce trial. Lastly, I want to express my gratitude to my friend Dan Delaney for proof reading this manuscript.

Mexico City, May 26th, 2013

Julián Romero Pérez

Declaration

I, Julián Romero-Pérez, declare that the thesis entitled: "The Use of Sound Intensity for Characterization of Reflected Energy in Small Rooms" and the work presented in it is my own. I confirm that:

• This thesis was written exclusively, or mainly while in candidature for a research degree at the University of Salford.

• Any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated, as I was previously doing a Ph.D. at the University of Huddersfield and later transferred to the University of Salford under the same main supervisor.

• Where I have consulted the published work of others, this is always clearly attributed in references.

• Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

• I have acknowledged all main sources of help in the Acknowledgements section.

• Where the thesis is based on work done by myself jointly with others, I have made

clear exactly what others did, and what I have contributed myself.

• Parts of this work have been published as conference papers, which are in chronological order:

1. Fazenda, B. M. & Romero-Perez, J. (2008). 3-dimensional Room Impulse Response Measurements in Critical Listening Spaces. Reproduced Sound 24: Immersive Audio, Apr 30, Brighton U.K. Proceedings of the Institute of Acoustics, 323-239.

2. Romero-Pérez, J. & Fazenda, B. M. (2009). Characterisation of the Directionality of Reflections in Small Room Acoustics. School of Computing and Engineering Researchers' Conference, University of Huddersfield, Dec 9th.

3. Romero-Perez, J., Fazenda, B. M. & Atmoko, H. (2009). Characterisation of Small Room Acoustics for Audio Production Reproduced Sound 25: The Audio Explosion, Nov 6, Brighton, U.K. Proceedings of the Institute of Acoustics.

This work was performed as a project funded by a few waiver at University of Salford from 2010 to 2013 and at the University of Huddersfield from 2008 to 2009. However, the views and analysis expressed in this Thesis are those of the author.

Please cite this document as follows:

Romero-Perez J. (2013). The Use of Sound Intensity for Characterization of Reflected Energy in Small Rooms. Ph.D. Thesis. Acoustics Research Centre School of Computing, Science and Engineering, University of Salford, Manchester, U.K.

Signed.....................................

Nomenclature

a) Abbreviations:

This list identifies some of the Abbreviations that are not necessarily defined every time they appear in the text.

CAD   computer-aided design.

CAM   computer-aided manufacturing.

CNC   computer numerical control.

DAW   digital audio workstation software (i.e. *Pro Tools*, *Logic Pro*,

                   Adobe Audition).

DirAC   Directional Audio Coding.

DOA   direction of arrival. $\{^{\circ}\}$

DSP   digital signal processing.

EDM   electro-discharge machine.

EDT   early decay time. $\{s\}$

ERB   equivalent rectangular band. {adimensional}

ERB_N equivalent rectangular bandwidth of the rectangular filters. {adimensional}

ETC   energy time curve.

dB   decibel (1/10 of a bel, i.e. 1B = 10 dB). $\{dB\}$

dBFS   full scale decibel reference to the maximum value of the signal s. {dBFS}

FIR   finite impulse response (Type of Filter).

FFT   fast Fourier transform.

Hz   Hertz {Hz = cycles/s}.

IFFT   inverse fast Fourier transform.

IID   interaural intensity difference. {dB}

ITD   interaural time differences (time delay). {ms}

JND   just noticeable difference.

LTI   linear time invariant filter.

MAA   minimum audible angle.

NFFT   number of points of the fast Fourier transform. {adimensional}

SPL   sound pressure level. {dB}

rms   root mean square value of a signal (s).

IR   impulse response. {Pa or dB}

IR_x impulse response in x-axis. {Pa or dB}

IR_y impulse response in y-axis. {Pa or dB}

IR_z impulse response in z-axis. {Pa or dB}

RIR   room impulse response. {Pa or dB}

RT   reverberation time. {s}

STFT   short-time Fourier transform.

SIRR	spatial impulse response rendering.
SNR	signal-to-noise-ratio . {dB}
SRP	steered response power.
TD	time domain.
TDOA	time difference of arrival. {ms}
TOA	time of arrival. {ms}
TDE	time delay estimation. {ms}
TOF	time of flight, also known as *TOA*. {ms}
VMC	vertical machine center.
1-D	one-dimensional.
2-D	two-dimensional.
3-D	three-dimensional.
3-D *IR*	three-dimensional impulse response.

b) Symbols:

This list identifies some of the symbols that are not necessarily defined every time they appear in the text.

$\mathbf{A}^{\mathrm{T}}$ transpose operator T for matrix or vector $\mathbf{A}$.

$\mathbf{A} = \begin{bmatrix} x,y,z \end{bmatrix}$ 3-D vector.

$\mathbf{A} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ transposed 3-D vector.

$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ matrix in expanded notation.

B   bel (unused unit of logarithmic scale).

c speed of sound as a function of temperature (T). {m/s}

c_0 speed of sound propagation in air @ $T = 20°$ C. {$c_0 = 343.4$ m/s}.

$\mathbf{C}^{*}$ complex conjugate operator applied to a complex number $\mathbf{C}$.

d distance between the microphones of the p-p intensity probe. {m}.

Δt_i time delay of the ith position. {ms}

$\Delta t_{dir-ref}$ time difference from direct to first reflection. {ms}

$\Delta t_{i \to i+n}$ time difference of adjacent position. {ms}

$\dfrac{\partial f(x)}{\partial x}$ partial derivative of a function of x.

E acoustic energy density. $\{$J/m^3 or (W s)/m^3 or kg/(m s^2) or Pa$\}$

E_{pot} potential acoustic energy. $\{$J/m^3 or (W s)/m^3 or kg/(m s^2) or Pa$\}$

E_{kin} kinetic acoustic energy. $\{$J/m^3 or (W s)/m^3 or kg/(m s^2) or Pa$\}$

$\overline{\varepsilon}_{\overline{\varepsilon}}$ overall mean error.

$\overline{\varepsilon}$ mean error.

$\varepsilon_{\%}$ percentage error. $\{\%\}$

$\left|\overline{\varepsilon}_{\%}\right|$ mean absolute percentage error ($MAPE$). $\{\%\}$

$\left|\overline{\varepsilon}_{\overline{\varepsilon}_{\%}}\right|$ overall mean absolute percentage error ($OMAPE$). $\{\%\}$

e exponential base, approximately 2.71828182 (*Matlab* function `exp`).

$f(x)$ function of independent variable x .

`fft`   fast Fourier transform function from *Matlab*.

f_c Critical frequency. $\{$Hz$\}$

$f_{\text{cut-off}}$ Cut-off frequency, corner frequency or break frequency. $\{$Hz$\}$

f_s Schroeder frequency. $\{$Hz$\}$

ϕ diameter of microphone capsules. $\{$mm$\}$

$\hat{\varphi}_i$ sound source i^{th} arrival estimated angle. $\{$ º $\}$

φ_i sound source i^{th} arrival real angle. $\{$ º $\}$

η pressure velocity correlation coefficient or radiation index. $\{$adimensional$\}$

ψ	diffuseness estimate proposed by Merimaa (2006). {adimensional}
$\hat{\psi}$	diffuseness estimate proposed by Ahonen (2009). {adimensional}
$h(t)$	room impulse response function in time domain. {Pa or dB}
$\mathbf{h}_{3D}(t)$	three-dimensional impulse response in time domain. {Pa or dB}
$H(\omega)$	transfer function in frequency domain. {Pa or dB}
$H[s(t)]$	Hilbert operator (real signal).
h_m	microphone height. {m}
h_s	source height. {m}
$\mathbf{h}_4$	Four channel orthogonal impulse response vector.
$\int f(x)dx$	integral of function of x respect to variable x.
$\hat{\mathbf{i}}$	x-axis component unity vector.
$\vec{i}$	sound intensity in time domain (real signal). $\{\,W/m^2\,\}$
$\vec{i}_a$	active instantaneous intensity in time domain (real signal). $\{\,W/m^2\,\}$
$\vec{i}_{complex}$	complex instantaneous intensity in time domain (complex signal). $\{\,W/m^2\,\}$
$\vec{i}_{active\ complex}$	active complex instantaneous intensity time domain (real signal) $\{\,W/m^2\,\}$
$\overline{\vec{i}}_{complex}$	envelope of complex instantaneous intensity.
$\overline{\vec{i}}$	envelope of intensity in time domain. (real signal) $\{\,W/m^2\,\}$

$\vec{i}_x$	envelope of intensity in x-axis (real signal). $\{\,W/m^2\,\}$
$\vec{i}_y$	envelope of intensity in y-axis (real signal). $\{\,W/m^2\,\}$
$\vec{i}_z$	envelope of intensity in z-axis (real signal). $\{\,W/m^2\,\}$
$\vec{i}_{inst}$	instantaneous intensity in time domain (real signal). $\{\,W/m^2\,\}$
$\vec{i}_{reactive}$	reactive intensity in time domain (imaginary part of $\vec{i}_{complex}$). $\{\,W/m^2\,\}$
$\left|\vec{i}\right|e^{\theta}$	sound intensity vector with amplitude and phase (Phasorial). $\{\,W/m^2\,\}$
$\vec{I}_a$	active instantaneous intensity in frequency domain (real signal). $\{\,W/m^2\,\}$
$\vec{I}_{complex}$	complex intensity in frequency domain (complex signal). $\{\,W/m^2\,\}$
$\vec{I}_{inst}$	instantaneous intensity in frequency domain (real signal). $\{\,W/m^2\,\}$
$\vec{I}_{reactive}$	reactive intensity in frequency domain (real signal). $\{\,W/m^2\,\}$
`imag`	imaginary part function in Matlab.
`ifft`	inverse fast Fourier transform function in Matlab.
j	complex operator also known as i, equal to $\sqrt{-1}=\left(0+j\right)$ $\{\,adimensional\}$
$\hat{\mathbf{j}}$	y-axis component unity vector.
k	a) discrete frequency index in a short Fourier transform. $\{Hz\}$
	b) wave number, angular wave number (Phase Coefficient) or
	propagation constant. $\{rad/m\}$

$\hat{\mathbf{k}}$ z-axis component unity vector.

l_{MFP} mean free path. {m}

L_x length of a room in x-axis. {m}

L_y length of a room in y-axis. {m}

L_z length of a room in z-axis. {m}

λ wavelength. {m}

m length of time window on a spectrogram. {samples}

`max`   maximum value function of a signal s in *Matlab*.

ms   milliseconds, i.e. 1×10^{-3}s {ms}

n discrete time index in a fast Fourier transform. {s}

N number of points. {adimensional}

p acoustic pressure (real signal). {Pa}

p_1 pressure of the microphone 1 of the p-p intensity probe (real signal). {Pa}

p_2 pressure of the microphone 2 of the p-p intensity probe (real signal). {Pa}

$\breve{p}$ analytic pressure signal (complex signal). {Pa}

p_+ a travelling wave in the forward direction (Positive axis). {Pa}

p_- a travelling wave in the backward direction (Negative axis). {Pa}

$\bar{p}_{inst}$ instantaneous average pressure between 2 microphones. {Pa}

$\tilde{p}$	complex acoustic pressure. {Pa}
P_n	n^{th} receiver position. {m}
$\mathbf{P}_4$	Quadraphonic momentum.
π	Pi, an irrational number approximately 3.141592 {adimensional}
Re	real part of a complex signal (real signal).
s	signal in time domain (real signal).
$\bar{s}$	envelope of a signal (real signal).
$\tilde{s}$	complex signal in time domain (complex signal).
$\tilde{S}$	complex signal in frequency domain (complex signal).
S	modulus, absolute value or magnitude of signal (real signal).
$\hat{s}$	Hilbert transform of a signal (s).
$\breve{s}$	analytic signal (s).
s_{real}	real part of a signal (s).
s_{imag}	imaginary part of a signal (s).
σ	standard deviation of a sample.
σ_ε	standard deviation error.
$\bar{\sigma}_\varepsilon$	mean of standard deviation error.
$\bar{\sigma}_{\bar{\varepsilon}_\%}$	weighted average of standard deviation of mean absolute percentage error. {%}

$\bar{\theta}$ Circular mean $\{\,^{\circ}\}$

θ_{m_i} i^{th} angle of arrival of first reflection. (i^{th} microphone position). $\{\,^{\circ}\}$

θ_{s_i} i^{th} angle of arrival of direct sound (i^{th} source position). $\{\,^{\circ}\}$

θ_{xy} angle sustained in x-y plane. $\{\,^{\circ}\}$

θ_{xz} angle sustained in x-z plane. $\{\,^{\circ}\}$

$r_{i_{direct}}$ i^{th} distance travelled from the source to the receiver. $\{m\}$

ρ_0 air density @ $T = 20^{\circ}$ C. $\{1.21 \text{ kg/m}^3\}$

T period. $\{s\}$

$T_{Celsius}$ temperature in Celcius degrees. $\{^{\circ} C\}$

t_i i^{th} time delay of direct sound. $\{ms\}$

t_{mix} mixing time. $\{s\}$

t_s centre time $\{s\}$

$\vec{u}$ particle velocity vector (real signal). $\{m/s\}$.

$\breve{\vec{u}}$ analytic particle velocity (complex signal). $\{m/s\}$.

$\tilde{u}$ complex particle velocity. $\{m/s\}$

$\hat{u}_0$ peak value of amplitude of complex particle velocity. $\{m/s\}$

V volume. $\{m^3\}$

ω	angular frequency. {rad/s}.
w	B-Format w signal in time domain (scalar). {Pa}.
w_{Hann}	Hann window.
W	B-Format W signal in frequency domain (scalar). {Pa}
W_{power}	acoustic power. {Acoustic Watts = W}
$\vec{x}$	B-Format x signal in time domain (vector). {m/s}
$\vec{X}$	B-Format X signal in frequency domain (vector). {m/s}
$X(\omega)$	input spectrum of a measurement system (stimulus). {m/s}
x_i	i^{th} microphone position distance in the x-axis. {m}
x_r	resultant sum of Cartesian components x-axis for an angle θ_i.
$\vec{y}$	B-Format y signal in time domain (vector). {m/s}
$\vec{Y}$	B-Format Y signal in frequency domain (vector). {m/s}
$Y(\omega)$	Output spectrum of a system (recorded signal). {m/s}
y_r	resultant sum of Cartesian components y-axis for an angle (θ_i).
$\vec{z}$	B-Format z signal in time domain. (vector). {m/s}
$\vec{Z}$	B-Format Z signal in frequency domain. (vector). {m/s}
$|\;|$	absolute value or magnitude or modulus of a complex or vector quantity.
$\|\;\|$	norm of a vector (Euclidean length of a matrix or vector).

$\langle\ \rangle$ time averaging.

$\{\ \}$ unit's separator or mathematical bracket.

$1°$ angular degree.

Abstract

The sound field in rooms of small dimensions used for music reproduction is characteristically different from that found in larger rooms for music performance such as auditoria. Key differences between small critical listening spaces and large auditoria are the vastly different ranges of energy decay, 100 ms for the former and up to 8 s for the latter, and its directional behaviour, typically non-diffuse for the former and approximating a diffuse field for the latter. Despite these substantial differences, most of the metrics developed to describe the sound field in large spaces are evoked to quantify the performance of small rooms. This project focuses on developing measurement methods to characterise temporal and spatial qualities of sound in small rooms.

A number of methods based on currently available acoustic probes have been developed. The implementation requisites and accuracy for each method has been quantified. Factors such as direction, time of arrival and strength of reflections have been extracted using signal analysis techniques based on the active instantaneous intensity and short-time Fourier transform. These factors are subsequently mapped to allow a description of their evolution through the energy decay in the room for a given measurement location. The best performing system, based on the use of one-dimensional p-p intensity probe mounted in a custom cradle, achieves a minimum overall mean error of 0.226 degrees and 2.971 degrees for the direct sound and first reflection respectively, which is near or below the measured human minimum audible angle (*MAA*). The method developed has direct applications in the quantification of small room acoustic sound fields for critical listening purposes.

Chapter 1 : Introduction

The characterisation of the early reflected energy is a topic that has been studied to improve the acoustic quality of spaces (Hyde, 1998). In the past decades, it has not been completely achieved by cause of the limitations of the measurement techniques and the measurement equipment available. However, in the last two decades, the development of powerful personal computers has helped to directly apply digital signal processing (*DSP*) techniques to obtain accurate impulse responses (*IR*) of the enclosures, which in the past were only implemented with dedicated analogue and digital measurement laboratory equipment. A complete characterisation of a sound field requires simultaneously quantifying the particle velocity ($\vec{u}$) in each Cartesian axis ($\vec{u}_x, \vec{u}_y, \vec{u}_z$) and the acoustic pressure (p). The development of reliable intensity probes has helped to improve this research because of its capability to map three-dimensional sound fields. Presently, there are still very few attempts to describe the capabilities of measurement systems to determine directionality of transient reflected energy for small rooms. Accordingly, that fact motivates this research (Toole, 2006, Toole, 2008).

The expression 'sound intensity' ($\vec{i}$) has been used in many fields of acoustics, such as the determination of sound power emission on steady state sources, characterisation of sound sources, sound fields, source location, source characterisation, measurement of sound absorption, acoustic impedance, sound power transmission through partitions, radiation efficiency of vibrating sources, active noise control of sources, acoustic holography, measurement of flow in non steady ducts and transient noise sources (Fahy, 1995).

The characterisation of early reflection fields has not been fully explored in terms of the effects of scattering and diffraction created in small rooms, since there is not a better angular resolution than ±3° using the three-dimensional p-p intensity probe. Accordingly, those mean errors ($\bar{\varepsilon}$) on the angle of early reflection arrivals (θ_{m_i}) are usually omitted for the sake of commercial strategy. Nevertheless, using a two-dimensional field

intensity probe consisting of four phase-matched measurement omni-directional microphones spaced on a cross shape, Elko reports that it may have an error less than 2 degrees when the source is a spherical wave and is emitting a single frequency of 1 kHz, also known as monochromatic frequency (Elko, 1984).

Despite the early attempts to characterise the directionality of direct sound and reflections in rooms, it was found that the characterisation of reflections is a challenge because of the limited number of separate early reflections, which can be safely analysed before the mixing time (t_{mix}) of the impulse response (*IR*). After passing the mixing time (t_{mix}), the high density of reflections overlaps during the reverberation stage and creates a difficult analysis for any reflection extraction. The diminishing signal-to-noise-ratio (SNR) of subsequent reflections is another issue related to the sensitivity of measurement using intensity probes. The motivation to carry on this work was determined under this present stage of insufficient spatial characterisation of early reflections in small rooms. Presently, there are many of acoustic models to assess an acoustic project, however, the options to characterise reflections in rooms using physical measurements is still not standardised.

1.1 The small room context

In the field of room acoustics, the enclosures are classified as two main types: large rooms and small rooms. This classification is mainly based on the room's volume and the behaviour of the room resonances across the spectrum. Several authors have tried to define the frontier between the two by using mathematical models (Schroeder and Kuttruff, 1962, Kuttruff, 1998, Skålevik, 2011). It was discovered that the relevant variables were the reverberation time (RT) and the room's volume (V). The objective is to determine the frequency where the statistical limit within a given sample of room resonances might be treated as a distribution. One of the most accepted models is the Schroeder frequency (f_s) limit formula (Schroeder and Kuttruff, 1962, Kuttruff, 2000), which was improved in accuracy later on in (Schroeder, 1987). Recently, Skålevik

Magne introduced another proposal, which still needs to be tested as a frequency transition crossover range between large and small rooms (Skålevik, 2011).

These big spaces are characterized by the critical frequency (f_c) also known as the Schroeder frequency limit (f_s), which imposes a minimum volume limit of 7,079 m^3 for a very wide range of music and 9,914 m^3 for speech (Davis and Patronis, 2006). Conversely, according to BS ISO 3382-2 (2008), the criteria for a small room stands at V < 300 m^3. According to Bolt and Roop, a typical example of small rooms have V < 70 m^3 (Bolt and Roop, 1950). Consequently, a room's volume is still a controversial parameter for defining a small room, due to the fact that there is not a contrasting boundary between small and big rooms, however, there is a smooth transition that complicates its classification.

A small room, in technical terms, can be defined by its volume size (V), additionally, in a more detailed analysis, may be defined by its density of reflections and duration of the sound decay. The principal characteristic is that their room impulse response (RIR) presents a sound field, which is dominated by the direct sound and early reflections, since its reverberation time (RT) is typically low at mid and high frequencies, even though it usually shows higher values at low frequencies. These small room dimensions tend to have a very low ceiling height in relation to its length and width. Accordingly, the main problem is that these spaces are non-diffuse and exhibit conditions where one of the planes has dissimilar distribution of reflections. This is independent of the type of acoustic treatment of the room. Another characteristic of small rooms is that they may have considerable absorption in some of the boundaries. The contents of the room can be objects that affect the scattering and absorption of sound, because in proportion, they represent a large area relative to the room. Its longest dimension is comparable to the wavelength (λ) of the lowest audible frequency i.e. 20 Hz (room dimension L < 17.17 m). One of its characteristics is that the sound field is not diffuse, since the lack of enough reflections coming from all directions, which is the lack of an isotropic sound field and lack of homogenous sound throughout the spatial coordinates (x, y, z). Typically, the propagation of a single wave may exhibit no more than 4 to 6 reflections before they reach inaudible levels. This is a result of the high absorption proportion

located in small rooms. The build up of room modes is noticeable at low frequencies; it generates an uneven sound decay, which concentrates energy in the low spectrum by repeated reflections in certain orientations, which create a non-diffuse sound field at the modal region. These Eigen frequencies tend to be sparse in the spectrum, creating large peaks and valleys with level differences up to 20 dB, which are perceived as coloration of sound in temporal decay and in certain frequencies (Bolt, 1939, Bolt and Roop, 1950). These room resonances, which are a function of frequency (f) and the room´s dimensions (L_x, L_y, L_z), reinforce and cancel each other at certain coordinates of space, creating nodes and antinodes. Antinodes present longer sound decay at some particular resonant frequencies, whereas nodes depict positions where certain frequencies cannot be heard thanks to its shorter decay. Small rooms tend to be very sensitive to their boundary conditions dictated by their acoustic impedance. Room boundaries tend to be reactive, therefore only the high frequencies are easily absorbed. There is a strong tendency to focus on the remaining low-frequency energy in the plane where the walls of the room exhibit less absorption. Small rooms do not exhibit a true reverberant field, since the reflection patterns do not have enough reflections to create the reverberant tail (Geddes, 1998) and this limitation on sample size prevents the use of statistical treatment such as a sample distribution. This is more noticeable at low frequencies, because small rooms with volumes less than 30 m^3 reveal several frequency spaced dominant axial, tangential and oblique modes, which are reinforced by the case of similar dimensions on length, width and height that lie in the same third octave band (Hopkins, 2007). Therefore, they are noticeably below the human perception limits, where the critical bands can be accurately approximated by third octave bands and the differences between two low frequencies below 500 Hz can be detected within a resolution of 1 Hz as a just noticeable difference (JND) (Fastl and Zwicker, 2007). Conversely, at low frequencies, the spatial perception is not accurate, as it is at mid frequencies. This is owing to the nature of perception of low frequencies, which is dominated by interaural time differences (ITD), while the perception of direction of mid and high frequencies is based on interaural intensity differences (IID) (Howard and Angus, 2001, Geddes and Lee, 2003a). The reverberation time (RT) in small rooms usually needs to be measured with a limited signal-to-noise-ratio (SNR) of 20 or 30 dB (RT_{20} or RT_{30}). This is by virtue of the fact

that sound decay is faster in a small room than in an acoustic large room. The sound energy is dissipated quicker and the perception of the room response and the reproduced sound cannot be separated as a result of the fact that the speaker's temporal response (ringing) may be in the same range of first reflections arrivals (Nelson, 1992). Therefore, in common small rooms it is very difficult to perceive a clear direct sound, which means absence of perception of an uncoloured direct sound component, unless the acoustic treatment approximates to an anechoic space (Geddes and Lee, 2003b). The critical distance is not a good descriptor because sources and receivers are so close that they are in a transitional field, which is located in between the acoustic nearfield and the reverberant field (Toole, 2008).

In order to make a difference between the characteristics of large rooms and small rooms, it is useful to describe the diffuse field theory. According to Davis (Davis and Patronis, 2006), if a room is big enough, it will present a high density of modes, even at low frequencies, since the space can be described by a statistical model also known as diffuse field. During the sound decay in the room, in these spaces there is an equal probability of sound waves arriving in any direction at any coordinate, by virtue of the fact that all the boundaries of the room have similar sound absorption coefficients (Barron, 1973). It relies on two assumptions: the sound energy is uniform in the field and the sound intensity ($\vec{i}$), which represents the flow of sound energy across time, is the same for all directions. The result is that there is no mean energy flow ($\vec{i}_a$) at any point in the room, because their directions are completely random. Consequently, the global effect of summing different sound intensity vectors ($\vec{i}$) cancel out. Since there is no noticeable concentration of sound pressure or any absence, the overall level is constant. Therefore, a diffuse field does not promote sound energy flow ($\vec{i}$) during the temporal decay, owing to pressure differences (Δp). The room boundaries within an idealised diffuse field should be perfectly reflecting. According to Kuttruff, the energy density (E) depicts the same value for all the possible coordinates of the room under steady state conditions (Kuttruff, 2000). Therefore, it is impossible to predict the directions of the sound particles by the diffuse wall reflections or by particles scattered by obstacles during their

free propagation. In mathematical terms, the time and ensemble averaging converges and are interchangeable. These are the characteristics of an ergodic process.

The energy density (E) of small rooms deploys a fast decay across time, i.e. its overall energy per unit volume vanishes before it can create modal activity. There is a lack of dense field reflections caused by absorption of mid and high frequencies and a poor or absence of dampening in the low frequencies. There is no mixing time (t_{mix}) to create a diffuse decay tail because the number of reflections is too small. Consequently, the sound field is neither homogeneous, nor statistical. Instead, there is a series of discrete early-reflected energy and an absence of a noticeable reverberation tail. Hence, the statistical theory, in which the Sabine equation is based, is not valid for small rooms (D'Antonio and Eger, 1986, Geddes, 1998, Kuttruff, 1998, Toole, 2008). The parameter called Schroeder frequency is still a useful descriptor for limiting a modal region in a room, regardless of its size (Davis and Patronis, 2006). Sound fields in small rooms are strongly influenced by the wave effects such as diffraction and interference at low frequencies, and there is no proper mixing of reverberation energy with the direct sound. The modal region shows separated resonances and reveals the small room size. The reverberation time (*RT*) is not a meaningful descriptor for small rooms, considering the fact that low frequencies, behave independently from mid and high frequencies and cannot be determined with statistical theory, nor with geometrical acoustics, on behalf of evidence of wave interference effects (Kuttruff, 1998).

Recalling the fact that small rooms do not have a valid application of the Sabine equation for reverberation time (*RT*); some reasons are mentioned: The low audible frequencies resonances are related to the phenomenon of matching wavelengths (λ) with any of the room dimensions (L_x, L_y, L_z). When a wavelength has half of the size of one dimension of these enclosures ($L = \lambda/2$), the first harmonic of a room mode appears by mixing the repeating reflections that build an axial mode, which is mentioned here because it is the strongest resonance in a room. Room modes may have n^{th} harmonics, but the most problematic appear below 250 Hz, because they are more spaced in the spectrum than the ones occurring at higher frequencies. The addition and cancelation of the consecutive

reflections in space are the components for the development of standing waves. The creation of standing waves promotes a longer decay of sound at the low resonating frequencies, and a progressively shorter decay of sound at mid and high frequencies, which do not sustain any room resonance. The assumptions of the applicability of the Sabine theory are based on the existence of a diffuse field. It is clear that small rooms cannot be diffuse because the contents of the room occupy a considerable volume of the room, interfere with the waves, and tend to diffract them. The density of reflections is smaller than in large rooms because of the considerable high frequency absorption effect, which is prominent in small rooms. It would be difficult to observe the diffuse decay tail in an impulse response (*IR*) due to its short nature. It tends to just show the direct sound and some early reflections. After that, the absorption of the small room dominates, and there is less chance to sustain a high density of reflections. Because of their small room size, they exhibit a short mean free path (l_{MFP}), which means that the sound waves will bounce more frequently, and consequently, will be attenuated by the absorption of the room boundaries in a faster way than in a large enclosure. Notwithstanding, the absorption in a large room is not an important factor because absorption is a limited resource for applying in such large areas. The increasing density of reflections creates a diffuse decay with several overlapping reflections, which tend to decay with an exponential tail.

1.2 Motivation

Recording studios have been typically implemented in acoustical terms in small rooms; the control rooms always have been its smallest working space among the recording spaces (Walker, 1996). To make this situation more extreme, commercial studios are becoming smaller because of the high cost of real estate properties and lack of opportunities to do big projects. These affordable digital audio workstations (*DAW*) record to any computer's hard disk and emulate hardware such as analogue music instruments and effect processors. They replace the expensive analogue recording equipment such as big consoles, effect processors and tape machines recorders. This paradigm change in the recording industry has introduced many problems concerning the

quality of the products that are recorded, mixed and mastered inside these poor acoustic spaces.

Additionally, there is an explosion of home music production activities such as recording, mixing and mastering, which usually take place in small enclosures, and are not ideal for such tasks. Small rooms usually have poor dampening at low frequencies, especially in the case of hard wall construction structures. This facilitates the formation of standing waves and non-diffuse sound decays. This condition decreases the immersion sense and the realism of the program to be critically assessed. Furthermore, the quality of sound is affected by the sound field interference caused by the relatively large size of the furniture and equipment in relation to the dimensions of the room (da Silva Vieira de Melo, 2002).

Coloration of sound perceived as a change of timbre, rhythm sensation and signal pitch is the main problem encountered in reproduced music within these spaces (Ando and Alrutz, 1984, Nelson, 1992, Halmrast, 2000, Toole, 2008). It appears in a complicated way, as it depends on temporal and spatial variables, which ultimately affect its frequency content in the form of comb filtering. In order to minimize these problems, acoustic treatment needs to be carefully applied to ameliorate the noticeable effects of strong modal activity and lack of precise stereo image. This common problem occurs because of improper treatment of early reflections at the sweet spot (Mäkivirta and Anet, 2001). The traditional ways for measuring rooms have not taken into account the importance of measuring the directionality of the sound energy flow ($\vec{i}$) and their effects in terms of perception of a precise stereo image. Therefore, a small room is only characterised completely when it is possible to assess the influence on the perception of early reflections, the sense of spaciousness, and ultimately, how the characteristics of neutral rooms are perceived (Geddes, 2009, Newell, 2012).

1.3 Objectives

This thesis focuses on the development of an accurate method to detect reflected energy using intensity in the framework of small rooms, where the accurate characterisation of early reflected energy becomes crucial to determine the proper room treatment for critical

listening spaces. Conversely, the acoustics of large rooms is appreciated as a mixture of its late reverberant field, and the directional and temporal distribution of early reflections, although the characterisation of early reflections is taken into account for the two cases. A different approach is taken in the design of both spaces. The large room designer's interest is focused on maximising the perception of clarity, intimacy, envelopment and loudness, (Cavanaugh et al., 2009, Griesinger, 2009). Conversely, in small rooms, the room designer tries to avoid cluttered early reflections that fuse perceptually with the direct sound and tend to blur the stereo image quality. The small room designs take advantage of the precedence effect while selecting proper acoustic treatment (Madsen, 1970, Haas, 1972, Walker, 1995, Hoeg et al., 1997, Begault et al., 2001, Geddes and Lee, 2003b, Begault et al., 2004).

The development of new measurement standards for acquiring three-dimensional impulse responses (3-D IR) is suggested. By combining multiple impulse responses (IR) in each Cartesian axis (IR_x, IR_y and IR_z), it is possible to characterise them in a more complete way than using just a single room's impulse response (RIR). This is because the acoustic pressure signal (p) lacks directional information, as it is a scalar unit. Conversely, the direction of sound may be determined by measuring the particle velocity ($\vec{u}$), and more precisely, by deriving the sound intensity ($\vec{i}$), which is obtained by combining acoustic pressure (p), a scalar quantity and the particle velocity ($\vec{u}$), a vectorial quantity. A suggested objective metric of acoustic quality in small spaces can be derived from it, and can be combined with the latest perceptual knowledge. Major emphasis is given to investigate whether using intensity probes could be plausible to obtain an angle resolution similar to or below the human minimum audible angle (MMA), or if it is needed to use more elaborated techniques. One of those techniques could be beamforming with its commercial solution such as the *Eigenmike*[1], which has been tested in detection of reflections in halls and in the cabin of a car in (Binelli et al., 2011, Farina et al., 2011). Another alternative technique may be higher-order spherical harmonics ($HOSF$) microphones (Daniel, 2003, Cotterell, 2009), spherical microphone arrays for capturing 3-D sound fields (Elko and Pong, 1997, Elko, 2000, Meyer and Elko, 2002, Huang and

---

[1] http://www.mhacoustics.com/mh_acoustics/Eigenmike_microphone_array.html
[Online accessed May 19th 2013]

Benesty, 2004), or a different transducer method based on an optical three-dimensional intensity probe (Cazzolato et al., 2005). The ultimate goal is to present a simple method that can be feasible to adapt it to current methods used in the field. The aims of this research are:

- To devise a system for characterisation of the acoustics of small rooms based on sound intensity probes.

- To test different types of intensity probes and compare them using objective metrics taken from room impulse responses (RIR) in order to characterise the acoustic quality of reflected energy based on its temporal behaviour and the shape of its spectrum.

The interim objectives to achieve this are set as:

- To implement and test two acoustic particle velocity ($\vec{u}$) probe systems, namely first order Ambisonic B-Format (*Soundfield* microphone) and the one-dimensional sound intensity face-to-face p-p probe mounted on custom developed cradle.

- To develop signal post-processing algorithms to be able to extract direction of arrival (DOA), signal strength and time of arrival (TOA) for reflection energy in small rooms using sound intensity ($\vec{i}$).

- To investigate objective acoustic measures for small room acoustic sound fields based on spatial measurements.

1.3.1 Post-processing and analysis

The project has been focused on the use of different types of intensity probes to acquire a three-dimensional impulse response (3-D IR) measured in small enclosures. The couple of probe types tested are: a) the *Soundfield* microphone, b) the face-to-face p-p intensity probe. The measurement of the acoustic energy arriving at the probe is detected as

instantaneous intensity vectors ($\vec{i}_{inst}$) that can show the evolution of the flow of reflected energy at any given time.

The information obtained can be used to extract temporal, spatial, and spectral information of the sound decay inside a room. The analysis is performed in the time domain (TD) and frequency domain ($STFT$) using instantaneous values of sound intensity ($\vec{i}_{inst}$) and energy density (E), which are used to compute the diffuseness estimate (ψ).

In the post-processing stage, several methods for measuring sound intensity ($\vec{i}$) are tested: the equations used to calculate the sound intensity compute the active complex instantaneous intensity ($\vec{i}_{\substack{active \\ complex}}$) and in parallel the instantaneous active intensity ($\vec{i}_a$). The calculation of the sound intensity changes slightly depending on the type of probe used.

In the case B-Format signals acquired with the *Soundfield* microphone, the instantaneous intensity ($\vec{i}_{inst}$) is calculated with the product of the acoustic pressure (W) and the particle velocity ($\vec{u} = X \cdot \hat{\mathbf{i}} + Y \cdot \hat{\mathbf{j}} + Z \cdot \hat{\mathbf{k}}$).

In the case of the one dimension p-p intensity probe rotated on each orthogonal axis, the instantaneous intensity vector ($\vec{i}_{inst}$) is found by calculating the instantaneous particle velocity ($\vec{u}_{inst}$), which is approximated by using the finite differences equation derived from the Euler equation (Fahy, 1995), (also known as momentum equation). Subsequently, by multiplying the calculated instantaneous particle velocity by the average instantaneous pressure ($\bar{p}_{inst}$) between the two microphones (for details please refer to page 41). From the analysis of these signals, a complete characterisation of the room response can be obtained to study such parameters.

The inputs to calculate the direction and time of arrival (DOA and TOA) of the early reflections are taken from these three simultaneous acquired room impulse responses (RIR). Another explored approach employs the time domain method (TD) by using the complex instantaneous intensity ($\vec{i}_{complex}$). By breaking the sound decay in variable time

window portions, it is possible to analyse the temporal evolution of the decay of the impulse response (IR). Afterwards, the active instantaneous intensity ($\vec{i}_a$), the complex instantaneous intensity ($\vec{i}_{complex}$) and the envelope of intensity are obtained ($\overline{i}$). Finally, relations among the correct detection of the direction of the early reflections and its perception as one of the factors considered in acoustical quality are investigated.

1.3.2 Measurement method description

A method based on using instantaneous active sound intensity ($\vec{i}_{act}$) obtained from B-format signals is adapted to capture the transient sound decay of small rooms. It attempts to map the spatial and temporal distribution of sound energy flow ($\vec{i}$) and the diffuseness estimate (ψ) in three dimensions and analyse the frequency content of early reflections to resemble the human capability to perceive them as echoes, image shift, tone coloration or cues to locate direction of the sound source (Ballou, 1991, Barron, 2009).

The arriving sound intensity at the receiver is carefully acquired with the highest temporal detail in order to preserve the impulse response of the room. The acoustic source used is the existing monitoring system of the control room under analysis. If the rooms to be measured do not have existing sound sources, then it is desirable to use an omni-directional one-point-source to simulate the natural sound propagation of real sound sources. These will address all the room modes and reflections that can be perceived when all the surfaces of the room are completely covered by the acoustic radiation of the omni-directional source. However, more importantly, it will simplify the determination of the directionality of reflections, since the source can be treated as a simple point in space instead of a distributed point source, which is complicated to characterise on any multi-driver monitor speaker.

An acoustic measurement system is implemented to generate a three-dimensional impulse response (3-D IR). The data is processed in order to calculate the direction and time of arrival (DOA and TOA) of the early reflections. Subsequently, the information obtained is further analysed in frequency bands by using the short-time Fourier transform method (

STFT) and the scale of the frequency is calculated with equivalent rectangular bands (*ERB*).

This analysis is motivated because of the circumstance that in practice, the direction of arrival (*DOA*) tends to differ across frequency by virtue of physical limitations of the microphone arrays limiting their accuracy within a certain frequency range. Moreover, the frequency content of early reflections varies depending on the level of dampening and its location in the room. It usually retains mid and low frequencies since high frequencies may be easily absorbed.

Conversely, the time of arrival (*TOA*) does not vary across frequency, because the front wave of a single reflection arrives at the same time for all the frequencies. Rooms usually exhibit steady state conditions, such as constant temperature (T) and no net flow, which guarantee a spherical sound propagation and a propagation of sound at a constant speed of sound (c_0). These conditions prevent refraction of the front waves. Another acoustic phenomena that can be taken into account is diffraction, which is sensitive to the size and contents of the room, and may affect the direction of arrival (*DOA*) of some of the early reflections depending on the position of the source and the receiver.

1.3.3 Defining objective characterisation for small room's acoustic fields

The characterisation of rooms needs to take into account that sound changes as a function of space and time and frequency. A room can be fully characterised by measuring the impulse response (*IR*). In the past, only one omni-directional measurement was enough for characterisation of the temporal response of a room, by the use of a pressure transducer. With the advent of the Fast Fourier transform (*FFT*) it was possible to analyse the frequency content of that room. In simple terms, an enclosure can be thought of as a linear time invariant (*LTI*) filter. A filter has a dual-domain property, because it can be treated as a custom frequency response with its custom phase response, and can be treated as a temporal decay characterised by its impulse response (*IR*)

Depending on the geometry of the room and the materials of its boundaries, the acoustic impedance of the boundaries may change the energy and the phase of the acoustic quality of the sound decay. The main characteristic of a real room is that its impulse response acts as a transient response. That is a combination of the arrival of the direct sound, the reflections that occur after it, and the late part comprises a complex blend of the exponential growth of reflections, which determine the reverberant part of the decay (Barron, 2009). This complex mixture of reflections varying in time and strength may determine the evolution and shape of the frequency content of the decay of sound, from the early reflections through the reverberant part.

1.3.3.1 Suggest objective metric of quality

Once a correct description of the sound decay is obtained, it is possible to continue with further studies on the perceptual effect of early reflection regarding the time and direction of arrival, strength, transient nature, and spectral characteristics. Any incoming sound is heard with different comb filters, which vary depending on where the location of the source was. Our brain analyses these comb filter differences, combining them with more information, such as the time of arrival (TOA), relative levels and directional filtering and with our personal shape of pinnae to discriminate the source direction of any incoming sound (Madsen, 1970, Ballou, 1991, Rubak and Johansen, 2003, Gentner et al., 2007). The effect of early reflections in our hearing system is subjective. One of its consequences is the increase of the spatial sensation, which can be explained under the Haas effect (Haas, 1972, Wallach et al., 1973, Barron and Marshall, 1981). Then it is possible to combine the objective spatial information with the state of art of psychoacoustic knowledge on early reflection perception to assess the acoustic quality of the measured rooms (Barron, 1971, Salomons, 1995). This knowledge states the conditions where some early reflections may be perceptually important or not relevant at any instant on the decay of the impulse response (IR).

The direction of the early reflections is an important factor in critical listening spaces, which are a special case of small rooms, given that location, direction and distances of objects are easier to distinguish than in larger rooms. This is because of the strong level

on direct sound relative to the closest attenuated early reflections that follow the direct sound. Early reflections bounce multiple times by traveling relatively short distances. Each time the sound decay encounters a boundary in the room, their initial level is progressively absorbed, and therefore it has a faster decay than in a big room. Additionally, a small room compared to a big room exhibits a shorter mean free path (l_{MFP}). There is a debate whether if the Hass effect is not relevant in these small rooms (Walker, 1995, Voetmann, 2007, Toole, 2008), because the majority of the sound field is held in the transitional field, not in the reverberant field. By cause of the typical sound absorption of rooms, the reflections tend to diminish the high frequency content and preserve the mid and low frequency content (Olive and Toole, 1989, Bech, 1995, Bech, 1996). If the early reflections arrive close enough to the direct sound, they are perceived as part of the direct sound with the aid of reinforcing their signal strength (Rettinger, 1968). If the early reflections possess the same phase, the reflection is coherent with the direct sound and the Haas effect may be working, however, in several cases the absorption of the room changes the coherence of the early reflections, which is detrimental on the quality of the perceived (Noxon, 1992).

1.4 Novel contributions

The two main contributions from the author of this thesis are:

1.- A measuring system using complex instantaneous intensity ($\vec{i}_{complex}$) for characterising the reflections in time domain (TD) method, and also applying it to the short-time Fourier transform ($STFT$) method for accurate measurement of reflected energy in small rooms.

2.- The design of a novel apparatus that enables spatial sound measurements, which uses a combination of concurrent laser-cross pointing technology. It enables accurate and consistent orthogonal axis rotation of a one-dimensional face-to-face p-p intensity probe.

1.5 Report layout

The report is organised as follows: This thesis is divided into nine chapters. Each one focuses on specific areas of the measurement system proposed.

Chapter 1 introduces the reader to the particularities of small rooms and their relevant features. It discuses the need of novel objective measures to describe these spaces. The motivation to accomplish this research is focused on small rooms and its directional and temporal attributes to improve the acoustics of critical listening rooms, such as control rooms. The objectives of the project are described focusing on the use of sound intensity ($\vec{i}$) and three-dimensional impulse responses ($3\text{-D } IR$) to characterise the temporal, spatial and spectral information of the decay of sound. Subsequently, a brief explanation of the post-processing and analysis of the data is presented, followed by the measurement system description. The objective measurements are described and linked to the suggestion of a subjective perception of reflections. The novel contributions are outlined at the end of the chapter.

Chapter 2 provides a literature review on areas such as temporal characterisation of sound decay by objective metrics, such as the ones based on the ratio of impulse response upon the signal integration. Subsequently, it mentions a review of the three-dimensional impulse response ($3\text{-D } IR$) techniques used for detection of reflections. The selected technique for this project is intensimetry applied to transient signals using instantaneous values of sound intensity ($\vec{i}_{inst}$) (Alfredson, 1980, Heyser, 1986a, Fahy, 1995). The next part focuses on early reflection measurements. Subsequently, measurements with *Soundfield* microphone, spaced-microphone arrays and p-p intensity probe are covered. A method using a rotated one-dimensional p-p intensity probe and a single microphone rotated in orthogonal axis is selected for experiments. The *Microflown* p-u probe is reviewed. At the end of the chapter, a table shows a comparison of the angular resolution of several measurement systems and the calculation method in chronological order.

Chapter 3 covers the applied theory and the calculation methods for the two probes chosen, (i.e. face-to-face p-p probe and *Soundfield* microphone). The concept of the four-

channel orthogonal impulse responses $\mathbf{P}_4(\vec{r},t)$ is the opening topic. In order to explain the basics for the advanced intensity quantities, an explanation of the following concepts is covered: the Hilbert transform $H\big[s(t)\big]$, the analytic signal ($\breve{s}$), and the envelope of the signal ($\overline{s}$) (Heyser, 1986a, Kuttruff, 2000, Johansson, 2008). Following this, the sound intensity theory is introduced. It is comprised of explanations of instantaneous intensity ($\vec{i}_{inst}$), active instantaneous intensity ($\vec{i}_a$) in time and frequency domain, complex intensity ($\vec{i}_{complex}$) and active complex instantaneous intensity ($\vec{i}_{active \atop complex}$). Lastly, the concept of diffuseness estimate (ψ) and equivalent rectangular bands (ERB) is explained. The use of complex intensity in time domain is a novel idea proposed by the author in order to improve the accuracy of estimation of incoming direction of early reflections. There are different approaches taken to calculate these quantities depending on the type of probe chosen, and the equations are depicted for each case. The next topic covered is the statistical treatment for the sample of angular measurement, which is comprised of the techniques used to minimise the uncertainty of the results. The parameters used are linear errors (ε) such as: the mean error ($\overline{\varepsilon}$), mean of standard deviation error ($\overline{\sigma}_\varepsilon$), the mean absolute percentage error $\left|\overline{\varepsilon}_{\%}\right|$ and the standard deviation of the mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$. Finally, the directional statistic parameters are defined. They consist of circular mean ($\overline{\theta}$) circular standard deviation (σ_o) and circular variance ($\sigma_o^{\;2}$).

Chapter 4 covers the acoustic measurement system by describing the probes-acquisition configuration on the two experiments undertaken, and secondly the description of the processing involved in the analysis of the data. The first topic talks about implementing the physical measurement system. The next section covers the early reflection real values calculation method used to obtain the angles of arrival of direct sound (θ_{s_i}) and the angle of arrival of first reflections (θ_{m_i}), which constitute the input and output variables used to perform a comparison with the measured ones. Subsequently, the time of flight (TOF) procedure is described. Thereafter, the Farina's exponential sine sweep method used to create a stimulus for room measurements is introduced (Farina, 2000). The following

topic is the extraction of the impulse response by direct deconvolution in the frequency domain. The probe configurations are the one-dimensional p-p probe with the custom-made cradle system and the *Soundfield* microphone used as an intensity probe. The accurate results presented in this thesis are achieved by using the custom-made cradle. Therefore, its design details are mentioned. The other custom aligner device is the one designed for the *Microflown* p-u probe. Subsequently, the *Soundfield* microphone used as an intensity probe is described. The next section is the processing and analysis of data where the time domain (TD) algorithm, the short-time Fourier transform ($STFT$), the peak detection algorithm and the envelope of a signal ($\bar{s}$) are explained.

Chapter 5 contains details of the Acoustic probe configurations, covering the measurement conditions and the repetitions needed to validate the accuracy of the method. The next section covered is the description of the measurement environment, which is comprised of: semi-anechoic chamber, monitor speaker influence, damping vibration on source and receiver, and the laser cross system to define spatial coordinates. It describes the laboratory test for single reflection measurements in the semi-anechoic chamber using the p-p probe and the *Soundfield* microphone. Afterwards, the single microphone multiple positions is mentioned. The next topic covered is the applications for the measurements in a control room where the application of a cheaper and more versatile *Soundfield* microphone for the measurement of a real room response with multiple reflections.

Chapter 6 describes the results of the measurements made in a laboratory test to quantify the accuracy of the estimation of direct sound and first reflections using the proposed system. The quantification is made by comparison of mean error estimations ($\bar{\varepsilon}$) and mean standard deviation of the mean errors ($\bar{\sigma}_{\bar{\varepsilon}}$) among the different techniques tested, such as mean absolute percentage error $\left|\bar{\varepsilon}_{\%}\right|$ ($MAPE$) or overall mean of a set many positions of samples P_i such as $\left|\bar{\varepsilon}_{\bar{\varepsilon}_{\%}}\right|$. The graphs showing the characterisation of the reflections are presented. The results of the angles of arrival of direct sound (θ_{s_i}) and first reflection (θ_{m_i}) are compared with a geometrical model. A couple of tables of

results are presented in order to compare results of different probes. Single reflection estimation graphs of the *Soundfield* microphone and the p-p intensity probe are presented. The following results presented are time domain (*TD*) method graphs. Afterwards, the results of estimation with instantaneous intensity ($\vec{i}_{inst}$) are compared with complex instantaneous intensity ($\vec{i}_{complex}$). Finally, the short-time Fourier transform method (*STFT*) results and the diffuseness estimate (ψ) graph are presented (Merimaa, 2006).

Chapter 7 presents the results of the applications of this system to a real room using the *Soundfield* microphone. A sketch of the room measured uses the right monitor speaker of a 5.1 channel control room. The short-time Fourier transform method is used to analyse the early reflections and the diffuseness estimate. Application of the measurement system is mainly to evaluate as feasibility study to demonstrate the measurement system can be applied (Romero-Pérez and Fazenda, 2009, Romero-Perez et al., 2009).

Chapter 8 contains the discussion of the experiments done in the laboratory test, and the results of the application to a real room. The accuracy and feasibility of these methods used is also discussed. The *Soundfield* microphone is suggested for practical measurements in real rooms because the p-p probe it is not practical and affordable for the common practitioner.

Chapter 9 presents a summary, the conclusions and further work derived from the previous chapters presented.
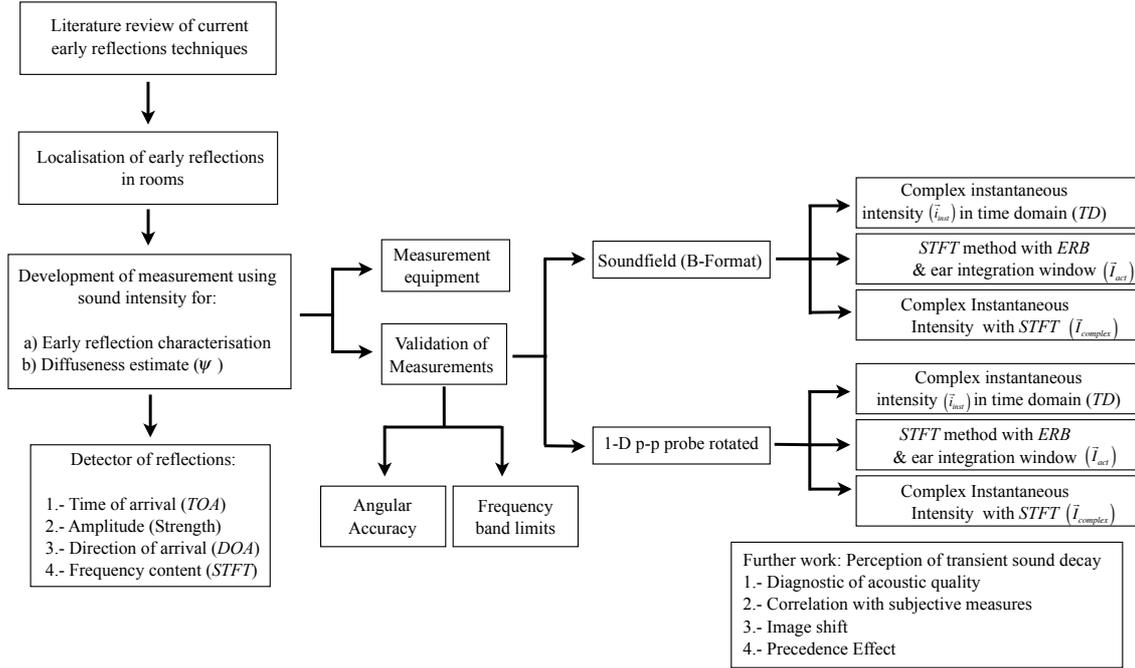
Figure 1.1: Schematic diagram of the approach taken in this research project.

1.6 Summary

This thesis aims to investigate the extraction and analysis of temporal and spatial distribution of early sound decay. A method based on using instantaneous active sound intensity ($\vec{i}_a$) obtained from B-format signals and a one-dimensional face-to-face p-p probe is adapted to measure small rooms. It attempts to map the spatial and temporal distribution of sound energy flow $\vec{i}(t)$ and diffuseness (ψ) acquiring a three-dimensional impulse response (3-D IR), and by performing an analysis of the frequency content of sound reflections. Special care was taken to obtain a measurement system with enough accuracy mapping the direction of reflections (θ_{m_i}) in order to resemble the human capability. Once the data is collected it is possible to extract the information in time and frequency domain and use it to infer issues of perception based on psychoacoustic models. The ultimate objective is to define useful descriptors to characterise the acoustic quality in critical listening rooms.

Chapter 2 : Literature review

The temporal characterization of sound decay is based on ratios of impulse response upon the total signal integration (Toole, 2006). Examples of these metrics are clarity index or early-to-late sound index (C_{80}), Definition or Deutlichkeit (D_{50}), early decay time (EDT), centre time (t_s), total sound level or strength (G). The early lateral energy fraction (LEF) uses a combination of an omni-directional and a figure-of-eight microphone (Barron and Marshall, 1981). The interaural cross-correlation coefficient (Damaske and Ando, 1972) $IACC$ uses a binaural dummy head microphone in order to compare two pressure signals. The listener envelopment LEV measures the perception of the level of surround of the sound that is coming from the reverberant field (Beranek, 2004).

It is now becoming more apparent that the total energy of the decay may be the key to achieve a sense of envelopment in small rooms (Beranek, 2008). The sound fields in small rooms need to be characterized with more detail in terms of lateral reflections and level of diffuseness, otherwise the results that may emerge from the mentioned measures still do not show clearly, which can be a desirable acoustic quality standard (Cox et al., 1993). The reason may be that they do not take into account such a fast process of sound decay and its spatial and temporal reflection's distribution.

The analysis of directionality of reflections is not new. The principal attempts have been made in controlled acoustic environments such as anechoic chambers but not for critical listening rooms, which usually are non-diffuse spaces because of their non-uniform absorption distribution. Microphone arrays have been used for mapping the energy using time delay differences instead of the coincident approach (Gover et al., 2004). Some developments to preserve the acoustical fingerprint of concert halls have been done using B-Format (Farina and Ayalon, 2003), which was an extension of a method proposed by Gerzon (Gerzon, 1975b). This method focuses on making complete characterisation of impulse responses. This approach may be the basis for developing further attempts to extract directional behaviour of decay.

Thiele and Meyer were probably the first to characterize directionality of reflections (Thiele, 1953, Thiele and Meyer, 1956). Yamazaki and Itow (Yamasaki and Itow, 1989), Sekiguchi, Kimura and Hanyuu (Sekiguchi et al., 1992), Okubo et al. (Okubo et al., 2001) and Gover, Ryan and Stinson (Gover et al., 2002) achieved the same outcome, but by using a time difference approach, which involves the use of omni-directional microphone arrays.

2.1 Measurement of early reflections using different probes

The simplest way to measure early reflections is to use a single pressure microphone. This approach has the advantage of being the most economical in terms of implementation. The downside is that it cannot give any spatial information about the reflections. It still can deliver some time of arrival (TOA) information and sound pressure level (SPL). Among all the research done in this way, it is worth noting that Moschioni (Moschioni, 2002) proposed correlation methods mixed with the envelope of intensity to detect the reflections by using a single pressure microphone, which delivers a one-dimensional impulse response.

In order to be able to completely measure early reflections, it is important to use a device that can map three-dimensional impulse responses of the rooms. There are three basic types of directional sound microphone techniques: a) separate microphone arrays, using discrete microphone arrays where the phase relationship of signals is not preserved correctly, b) coincident microphone arrays, which minimize the distance between the capsules in order to converge to a point in the space, and c) intensity probe configurations (face-to-face p-p intensity probe), which use a pair of measurement-grade pressure calibrated capsules placed at fixed distances (Fahy, 1995). The signals are processed to obtain the sound intensity vector ($\vec{i}$) along the axis where the probe is aligned to measure the estimations of direction of arrival (DOA) by calculation of the angle of arrival of direct sound (θ_{s_i}) and its first reflections (θ_{m_i}) (Abdou, 1994).

2.1.1 *Soundfield* microphone

Coincident microphones have been used to capture the sound field at a single point in space to preserve the same phase and level arriving at the capsules. This ensures a coherent mixture of the incoming signals, which are free from phase cancellations. In practice, there is a constraint of a minimum distance achievable between both microphone capsules directly related to its size. In 1933, Blumlein (Blumlein, 1933) proposed a stereo recording technique known as XY configuration by using a pair of ribbon microphones also known as velocity microphones. Each microphone had a figure-of-eight directivity. This array was designed to capture an accurate stereo image, which usually is aligned to a single plane in space (i.e. *x*-*y* plane).

In 1975, Gerzon published for the first time the design details of a new three-dimensional microphone array consisting of a tetrahedral array of microphone capsules and described a spacing compensation performed by electronic circuits to ensure successful coincidence of all outputs at high frequencies where diffraction effects degrade the localisation of sources (Gerzon). By 1977, a patent of this device was published by Craven & Gerzon (Craven and Gerzon, 1977). They explain that they developed a generalization of the Blumlein pair and developed a three-dimensional, sound-capturing microphone. Nowadays, it is known as the *Soundfield* microphone. It captures four signals from closely spaced capsules located on the edges of the faces of a tetrahedron. The raw data signals are taken from studio grade capsules, which have sub-cardioid directivities. In order to obtain a coincident microphone, some mathematical manipulation is applied to the four signals. The purpose is to create a virtual microphone, which consists of three virtual figure-of-eight directivity patterns that are overlapped at the origin, and are aligned on each orthogonal axis (*x*, *y*, *z*).

On the other hand, the accuracy of the localization of sound is compromised by physical and engineering limitations. The minimum distance between the capsules in a tetrahedral array still depends on the chosen diameter of the microphone capsules. Consequently, this determines the size of the array. It limits the safer frequency region to work without distortion of the sound field. This effect is caused by diffraction, which happens when

the wavelength (λ) of the sound has similar dimensions as the diameter of the capsules (ϕ). The size of the capsules also implies an inevitable variation of the directivity patterns across frequency, which is later compensated with some filtering. In the beginning, it was implemented with analogue electronic solutions. Currently, it is performed with digital filters implemented in a dedicated hardware, and recently, the filtering may be prepared offline with software.

At the University of Salford, the available commercial solution that can be used for the objective of detection of reflections is the *Soundfield* ST350 microphone. This portable version claims that the B-format signal is derived at 'single point', which lies inside the tetrahedral array. The accuracy of this microphone comes from the concept that its four signals are completely phase coherent. This is achieved by applying some analogue corrective equalisation to the spectrum of four signals created by a tetrahedral array of sub-cardioid capsules.

Humphrey did a study in 2006 of the angular accuracy of the direct sound using a source speaker and the original *Soundfield* microphone(Humphrey, 2006). A company called Calrec originally developed this microphone. Afterwards, the *Soundfield* company took over its manufacturing and distribution. The Calrec microphone model used in this study is the CM4050. It was used to calculate the estimation of the angle of the direct sound (θ_{s_i}) using the estimated particle velocities (X, Y, Z) obtained from the recorded B-Format signals. Humphrey reports a horizontal plane angular accuracy of ±1°, for the direct sound. He did not report any measurements for any early reflections, since the nature of his research was focused on location of sound sources only.

Essert (Essert, 1997) developed a method based on the *Soundfield* microphone to measure the proposed lateral energy fraction (LEF), the instantaneous lateral fraction(LEF_{inst}), and the directional fraction (DF) to describe the evolution of the sound field for hall design. However, since the use of the *Soundfield* microphone is not totally accepted as a measurement-grade device because of the high variability on the figure-of-eight directivity pattern, which tends to vary across frequency specifically above 5 kHz and also across the angle of incoming sound because of the combination of the distorted

figure-of-eight directivity patterns involved in the spatial measurement. This may be the technical reason why this method has not been popularised at least using the *Soundfield* microphone.

Dimoulas et al. (Dimoulas et al., 2007) (Dimoulas et al., 2009) utilised a B-Format hybrid approach to estimate the exact location of a source by using two *Soundfield* microphones. This was achieved with a combination of coincident microphones and the delay difference technique of triangulation of the source. However, this approach was used only for direct sound, not for early reflections. The most recent study of the location of sources using the *Soundfield* ST 350 microphone and intensimetry is found in Wierzbicki et al. (Wierzbicki et al., 2013). He makes a detailed study of the usable frequency range for the determination of direction, which is reported to be from 125 Hz up to 4 kHz. The preference for using a *Soundfield* microphone over any intensity probe because is user-friendly, practical and more cost-efficient than any p-p or p-u intensity probe.

A more complete study using the *Soundfield* microphone for detection of direct sound and early reflections was presented by Günel et al. (Günel et al., 2005) (Günel et al., 2007). She tried to implement the *Soundfield* microphone with a higher resolution. The B-Format was processed performing a wavelet packet decomposition method across a horizontal *x-y* plane with 360º. She shows that the range of biggest errors in the azimuthal plane was between angles of 135º and 240º, where errors have values between 10º and 25º. The mean error ($\bar{\varepsilon}$) reported for the estimation of source direction is $\bar{\varepsilon} = \pm 7°$; it implies that the same resolution for the early reflections must be the same as the resolution of the estimation of the direct sound.

However, experience of the author performing measurements of early reflections indicates that this is not a correct assumption. Measurements of error on reflection estimation have to be treated separately, considering the degree of scattering effect that happens in any reflection. This introduces some uncorrelated reflections with the direct sound, which makes the analysis of reflections difficult to achieve without prior knowledge of the direction of such reflection. Additionally, the use of a directional sound source with more than one drive also creates different arrival times of the reflections

because of the difference of time response of the different drivers (Romero-Perez, 2010). Fazenda and Romero-Perez (Fazenda and Romero-Perez, 2008) used B-Format signals to extract directionality of sources using particle velocity ($\vec{u}$) only. However, poor localization was found at angles near 45º when the instantaneous particle velocity vector was calculated in the time domain. Consequently, some of the results of this research were not included in the body of this thesis (see **Figure 7.3**). Following Merimaa's proposal using the spatial impulse response rendering (*SIRR*) (Merimaa and Pulkki, 2005), Romero-Perez and Fazenda explored the sound intensity method applied to the *Soundfield* microphone applied to small rooms, by applying it to a control room (see **Figure 5.13** and **Figure 7.1**) (Romero-Pérez and Fazenda, 2009, Romero-Perez et al., 2009).

It should also be noted that Farina (Farina, 2007), Faller (Faller and Kolundzija, 2009), Spenceley, and Wiggins (Spenceley and Wiggins, 2009) reported a limitation of the *Soundfield* microphone for accurate estimation of reflections with deviations up to 10º. This is because of the lack of uniform first order polar-patterns across the entire audible range. The asymmetries and distortion of the original directivity patterns are present since the there is a minimum distance between the microphone capsules. Above 12 kHz phase errors are present, the polar directivity of the figure-of-eight and the omni-directional patterns are distorted as a result of the first order spherical harmonic resolution, and also because of the limitation of the correction filters first developed by Gerzon (Gerzon) and later implemented by the company Calrec, which finally were improved by the *Soundfield* company. Recently, Heller and Benjamin tried a new philosophy for the design of the filters. They have determined the diffuse field as the best solution to calibrate a *Soundfield* microphone with flat diffuse field target (Heller and Benjamin, 2012), this will improve the future of this device as a potential measurement-grade microphone.

Recently, a new technique to display the reflection's pattern by using a *Soundfield* microphone as a three-dimensional hedgehog shape was presented by Bassuet (Bassuet, 2010). He proposes new spatial indicators that analyse the direction of the sound energy. Specifically, the spatial decomposition metrics are: *LH* (ratio of Low lateral and High

lateral energy) and *FR* (ratio of Front lateral versus Rear lateral energy). These parameters need to be quantified with higher accuracy than the spatial resolution of the *Soundfield* microphone in order to be assessed perceptually without biases.

On the development of a better spatial resolution, Craven et al. (Craven et al., 2010) have patented several sets of higher-order Ambisonic tangential array of microphones. Since the new proposed arrays are more compact than the current *Soundfield* microphone, they seem to be a good option, which have better directivity patterns at higher frequencies than the current *Soundfield* microphone. Benjamin (Benjamin) has tested some of these proposals with real prototypes. One major concern is that more capsules are needed to cover the higher-order spherical harmonics. Currently, there are not commercial solutions available for higher-order ambisonic microphones. However, there are other patents interested in creating a better version of the original *Soundfield* microphone by introducing a higher-order version called "Sound Field microphone" proposed for the AKG Acoustics GmbH company (Friedrich, 2012). It will request an increased number of channels, and calibration of the capsules will be crucial to obtain a good result from it. In general, this solution is more expensive in terms of resources and materials. Recently, another way to measure sound intensity ($\vec{i}$) using optical sensors was proposed by Cazzolato (Cazzolato et al., 2005), and may be useful for some fields of research of transient sound decays.

2.1.2 Spaced microphone beamforming and directional arrays

Several researchers have used different geometries of spaced-microphone arrays to locate the direction of the incoming sound. The majority of them are only interested in the location of the direct sound. The geometric model is usually assumed. The sound propagation inside a room is treated as rays and the virtual sources are modelled to create the reflections. This assumption is valid when the wavelength (λ) of the sound is smaller than the dimensions of the reflecting wall. Yamasaki and Itow (Yamasaki and Itow, 1989) used a four-omni-directional microphone array using a single microphone located at the origin of the three-dimensional axis and later it has been reviewed in (Havelock et al., 2008, Kuttruff, 2009). The remaining three microphones were located on each axis at

a close distance of 3-5 cm. The method of analysis was cross-correlation and intensity. Roberts reviewed the tetrahedral array and compared time of arrival (TOA) and beamforming methods also using an octagon microphone array (Roberts, 2009). Later, Sekiguchi used four-omni-directional microphone array located at each of the four apexes found in a regular tetrahedral geometry. The Japanese company Ono Sokki has developed a 3-D intensity probe measurement system Model MI-6420 "Tetra-phone" based on this idea[2]. By extracting the impulse response (IR) from each microphone position, the analysis was carried on in time domain (Sekiguchi et al., 1992). Choi et al. improved Sekiguchi's tetrahedral array by adding an additional centre microphone in order to have a common time of arrival reference (TOA) for the unlikely case of simultaneously arriving impulses. This solution improves the comparison between 5 microphones to solve the problem of equal time impulse arrivals (Choi et al., Choi et al., 2003). Later Rechenberger tested this configuration and implemented minimum phase compensation, a TOA algorithm and up-sampling method to claim an astonishing 2º angular accuracy. Nevertheless, this is theoretical, because the practical limitations of sample frequency and positioning accuracy of the microphone array were addressed and there was no reported method for measuring the accuracy of the system (Rechenberger, 2009). Cazzolato measured the energy density with a 3-D microphone array consisting of 4 microphones in a tetrahedral array and also measured the 6 and 7 microphone configuration analysing the error thoroughly, and concluding that the 4 microphone array is enough for measuring energy density (Cazzolato, 1999). However, Miah (Miah, 2009) (Miah and Hixon, 2010) has done the most complete design project of a three-dimensional broadband intensity probe, comprising 7 microphones arranged in a two concentric array along three axis, which are 120º apart. His new design of intensity probe can simultaneously cover two frequency ranges at the same time 200 Hz – 1 kHz with the low-frequency array and 1 kHz - 6.5 kHz. The angular accuracy reported was 4º ± 3º using a finite difference intensity method.

Another more powerful method used for location of reflected energy is beamforming. Gover used this approach after finding that the latter approaches lack of capability to resolve the arrival of simultaneous reflections (Gover, 2001). He developed an acoustical

---

[2] http://www.onosokki.co.jp/English/hp_e/products/keisoku/s_v/mi6420.html [Online accessed on May 22nd 2013]

transparent array with 32 microphones. By using beamforming techniques, a spatial resolution of 22º was successfully achieved. This technique is useful for differentiating simultaneous arrival of reflections because it can focus the beam in different directions independently of time of arrival (TOA). Ideally, this technique is the most robust for determining early reflections and simultaneous reflections, which occur in the reverberant tail. The downside is that in order to obtain a reasonable finer angular resolution, the number of microphones needs to be considerably larger. Hence, the measurement system needs to be more expensive in terms of equipment and computational power.

The work of Merimaa (Merimaa et al., 2001), Peltonen (Peltonen et al., 2001) and Merimaa's thesis (Merimaa, 2006) set the standards of extraction of directional impulse responses for the characterisation of reflections. It was achieved by displaying time, frequency and diffuseness in a single graph based on a spectrogram. Its measurement system consisted of two microphone arrays: a *Soundfield* microphone, and a custom-made 3-D p-p intensity probe denoted as the TKK-3-D microphone array, which has two concentric three-dimensional intensity probes. The inner one uses six electrets microphones in a configuration of three orthogonal microphones pairs with a distance of 10 mm between each microphone aligned in each orthogonal axis. The outer one uses a similar configuration with a distance between the microphones of 100 mm.

2.1.3 p-p type probe

Probably one of the first attempts to measure sound intensity on noise sources using a pair of phase-matched microphones is documented in a B.Sc. Thesis made by Sague (Sage, 1975) and supervised by Fahy in 1975 from the University of Southampton U.K.

According to (Fahy, 1977, Chung, 1978, Fahy, 1995) it was not until 1976, when the introduction of the p-p probe intensity measurements combined with an *FFT* analyser was practically adapted to real measurements of sound intensity ($\vec{i}$) as it is known in its modern formulation. The use of a single dimension p-p intensity probe for 3-D *IR* assumes time invariance is made, where the superposition of the ideal directivities is performed with different directivity of microphones located in the same position at

different times. A review of relevant works on numerical implementation of the intensity using different microphone techniques follows.

The first investigation of the transient intensity was performed by Alfredson (Alfredson, 1980). He employed two face-to-face condenser microphones without a spacer to measure impact noise by applying the instantaneous intensity formula in the time domain, and comparing it with the intensity calculated for plane waves. Modern techniques for measuring intensity with p-p intensity probe use the face-to-face configuration and a plastic solid spacer to avoid undesirable diffraction effects, and increase the usable frequency range an octave above the finite difference error (Fahy, 1995, Jacobsen et al., 1996).

2.1.3.1 One-dimensional p-p intensity probe rotated on orthogonal axis

The nature of sound decay, inherently, is a transient signal. For its proper analysis, Abdou (Abdou, 1994) recommends the measurement of active instantaneous intensity ($\vec{i}_a$) (Fahy, 1995) to locate the direction and strength of the direct sound and the subsequent reflections. This is accomplished by using a one-dimensional face-to-face p-p intensity probe rotated on an orthogonal axis.

2.1.3.2 Single microphone rotated on orthogonal axis

Omoto (Omoto and Uchida, 2004) and Fukushima (Fukushima et al., 2006) used a single microphone for measuring intensity in a way that the limitations of phase matching between the microphone pair is not an issue. The downside is that the procedure of measuring is too time consuming. All the measurements need to be done individually, and time variance is a real concern. Another problem can be the errors in spatial position induced by the use of this technique.

Omoto and Uchida (Omoto and Uchida, 2004) used the active instantaneous intensity ($\vec{i}_a$) approach, although calculated the envelope of the instantaneous intensity ($\overline{i}_{inst}$) to detect

the direction of the reflection as the combination of these peaks of envelopes in orthogonal axis.

Fukushima et al. (Fukushima et al., 2006) implemented a measurement system of sound intensity using a single microphone. The microphone was measured in four different positions of an array of microphones similar to the Yamasaki and Itow (Yamasaki and Itow, 1989) approach. The sound intensity ($\vec{i}$) was measured by taking subsequent measurements, changing the position of a single microphone to the places where the array had a microphone. This approach makes it possible to overcome the problem of phase matching, which adversely affects the accuracy of the reflection's angle (θ_{m_i}) and its strength using a p-p intensity probe. The downside is that it requires discrete measurements, which may not be time invariant. Another downside is that is very time consuming to take a single compound measurement, which relies on the accuracy of placement of several positions of the microphone.

2.2 Other methods

2.2.1 *Microflown* p-u probe

Further advances have been reported with a new approach based on hot wire transducers, which are sensitive to temperature changes. These are known as particle velocity transducers (Fahy, 1995), which are related to the fluctuations of value of particle velocity at that given coordinate. The p-u intensity probe was invented in 1996 by Hans Elias de Bree in (de Bree, 1997) and later commercialised in (de Bree, 2003). The company that manufactures this probe is called Microflown Technologies and is based in The Netherlands. This solution drastically minimizes the distance between the transducers, and is therefore, likely the smaller coincident array, showing the full bandwidth across frequency with a single probe. Owing to its small size, it is also possible to use the p-u intensity probe in smaller places than the p-p intensity probe (Druyvesteyn and de Bree, 1998).

In 2002, an American company named Meyer Sound[3] created the first measurement system for measuring 3-Dimensional impulse responses of rooms using a custom p-u microphone array. It consisted of one pressure microphone and three orthogonal particle velocities probes from *Microflown* (model PU regular one-dimensional) and custom hardware and software. The acoustical parameters such as *RT, EDT,* etc., were calculated using the Euclidean norm value of the instantaneous sound intensity ($\left\|\vec{i}_{inst}\right\|$) instead of the squared acoustic pressure (p^2). This process delivers lower floor noise than using only acoustic pressure, since the acoustic pressure (p) and particle velocity ($\vec{u}$) are uncorrelated signals and each transducer type senses them separately. When both are multiplied, the net effect is to obtain a cleaner signal. It has passed through a kind of filter, which removes unwanted noise by making it smaller after multiplying two small noise signals. Identification of the source of reflections is possible by analysing the polarity of each of the Cartesian axis of each early reflection by plotting them as a traditional Reflectogram, where the squared of the pressure signal (p^2) is converted to a dBA level (2002).

According to Raangs, the three-dimensional *Microflown* p-u intensity probe model USP (Ultimate Sound Probe) was introduced to the market (Raangs, 2005), with its signal conditioner model MFSC-4. It is reported that it has an accuracy of 7º in the estimation of the particle velocity ($\vec{u}$) when using steady state noise sources (Yntema et al., 2006). Nevertheless, the accuracy for detection of early reflections has yet to be fully reported. The aim of this project is to select a measurement-grade instrument that can surpass the minimum audible angle resolution of the human hearing system. Yntema, in his Ph.D. thesis, (Yntema, 2008) still finds inherent problems on the self-noise of a p-u *Microflown* intensity probe and on the angular accuracy achieved, but stresses the advantage of using only particle velocity sensors in order to overcome the discrepancies on signal-to-noise-ratio (SNR) found using pressure and particle velocities together. There was a low-frequency accuracy problem in the probe given that there is thermal noise in the hot wires addressed by J. W. van Honschoten in (van Honschoten, 2004). A new *Microflown* probe

---

[3] http://www.microflown.com/files/media/library/Applicationnotes/meyersound_3dp.pdf
[Online accessed on April 10th 2013]

has been proposed with four pairs of wires instead of two. It exhibits a lower inner noise level because of an increased sensitivity than the one found in a previous two-wire design (Yntema and van Honschoten, 2010). Nevertheless, the application of the *Microflown* probe for measuring early-reflected energy is challenging because the requirements for analysis of transient noise sources are more demanding than the requirements of steady state noise sources.

Author	Year	Measurement System	Angular resolution	Calculation method
Elko (Elko, 1984).	1984	2-D p-p intensity probe.	2º @ 1 kHz	Finite differences intensity method.
Sekiguchi (Sekiguchi et al., 1992).	1992	4 omni-directional microphones tetrahedral array.	3º	Deconvolution technique with tone burst stimuli.
Abdou (Abdou, 1994).	1994	1-D p-p intensity probe rotated.	3º@ 250 Hz	Instantaneous intensity ($\vec{i}_{inst}$) in time domain.
Van Lancker (van Lancker, 2000).	2000	8 omni-directional microphones in the corners of a cube.	1º or 2º	*TDE* method.
Gover (Gover, 2001).	2001	32 omni-directional microphones spherical array.	22º	Beamforming.
Günel (Günel et al., 2005, Günel et al., 2007).	2007	*Soundfield* microphone.	7º	Wavelet packet Decomposition.
Yntema (Yntema et al., 2006, Basten et al., 2009).	2006 2009	3-D *Microflown* p-u probe.	a) 7º b) 10º with no correction & 3.5º corrected	a) particle velocity ($\vec{u}$) only. b) instantaneous intensity ($\vec{i}_{inst}$) method .
Rechenberger (Rechenberger, 2009).	2009	5 omni-directional microphones tetrahedral array.	2º	*TOA* algorithm with up-sampling and minimum phase *IR*.
Miah (Miah, 2009) (Miah and Hixon, 2010).	2009	7 omni-directional microphone array.	4º ± 3º	Finite Differences Intensity method.
Tervo (Tervo, 2010, Tervo, 2012).	2012	3-D G.R.A.S. p-p intensity probe and TKK 3-D probe.	1º	*SRP* method combining *TOA* and *TDOA* algorithm.
Romero-Perez (Romero-Perez, 2011, Romero-Pérez, 2011).	2013	1-D p-p intensity probe rotated with custom cradle.	2.971º ±0.226º	*STFT* method using instantaneous intensity ($\vec{I}_{act}$) and Circular statistics

Table 2.1: Comparison of early reflections measurement systems, which reported quantifiable data.

2.3 Summary

After the survey on measurement of early reflections, it was found that a formal investigation on the accuracy of estimation of reflections was needed. In order to improve the B-Format angle estimation a new approach was taken. Therefore, the motivation of using active instantaneous intensity ($\vec{i}_a$), complex instantaneous intensity ($\vec{i}_{complex}$), envelope of intensity ($\bar{i}$) and short-time Fourier transform ($STFT$) approach was implemented in the present work.

Chapter 3 : Applied Theory

The implementation of a measurement system is based in intensimetry. The concepts covered are the acoustic energetic quantities used for the diffuseness estimate, envelope and analytic function used to obtain the peak detector in the post-processing analysis, and several topics covering sound intensity such as the Euler equation to approximate the particle velocity ($\vec{u}_{inst}$), instantaneous intensity ($\vec{i}_{inst}$) and active instantaneous intensity ($\vec{i}_a$).

The second part covers the post-processing equations used, such as the Exponential sine sweep, the deconvolution equation to obtain the impulse response of an acoustic measurement, and the application of the single vector/matrix that represents the four-channel orthogonal impulse responses.

3.1 Sound Energy analysis

3.1.1 Four channel orthogonal impulse responses

The propagation of an acoustic wave in air is best described with a single function (potential) that depends on the spatial coordinates x, y and z and the temporal variable t (Heyser, 1986a) and (Stanzial and Prodi, 1997). This idea has been patented in (Stanzial et al., 2005). The main treatment of energetic analysis used for calculation of sound intensity ($\vec{i}$) is based in a generalisation of quadraphonic impulse responses. According to the acoustic quadraphony theory, it is possible to completely characterise a sound field using the following definition of quadraphonic momentum ($\mathbf{P}_4$) and the 3-D space vector ($\vec{r}$) (Stanzial et al., 2000, Stanzial et al., 2002, Stanzial et al., 2005, Bonsi et al., 2007):

$$\mathbf{P}_4(\vec{r},t) = \left[\begin{array}{cccc} \dfrac{p(\vec{r},t)}{c} & \rho_0\vec{u}_x(\vec{r},t) & \rho_0\vec{u}_y(\vec{r},t) & \rho_0\vec{u}_z(\vec{r},t) \end{array} \right] \qquad (3.1)$$

Where $\vec{r}(x,y,z)$ is the 3-D Euclidian space vector.

In the case of B-Format signals, there is a similar analogy for implementation of vector that represents a four-channel orthogonal impulse response, which can describe the sound field at any fixed coordinate by the dual-domain short-time Fourier transform method:

$$\mathbf{h}_4(\omega,t) = \left[\quad W(\omega,t) \quad \frac{\vec{X}}{\sqrt{2}}(\omega,t) \quad \frac{\vec{Y}}{\sqrt{2}}(\omega,t) \quad \frac{\vec{Z}}{\sqrt{2}}(\omega,t) \quad \right] \qquad (3.2)$$

The B-Format W, $\vec{X}$, $\vec{Y}$ and $\vec{Z}$ signals are obtained by extracting the impulse response respectively (please refer to section 4.2.4). For a non B-Format signal, just use the following not scaled expression:

$$\mathbf{h}_4(\omega,t) = \left[\quad p(\omega,t) \quad \vec{u}_x(\omega,t) \quad \vec{u}_y(\omega,t) \quad \vec{u}_z(\omega,t) \quad \right] \qquad (3.3)$$

3.1.2 Hilbert transform $H\big[s(t)\big]$

It is widely known as a mathematic transform, however in reality it is just a mathematic operator (Thrane et al., 1984, Heyser, 1986a, Kuttruff, 2000, Johansson, 2008) defined as:

$$\hat{s}(t) = H\big[s(t)\big] = \frac{1}{\pi} \int_{-\infty}^{+\infty} \left(\frac{s(t-t')}{t} \right) dt' \qquad (3.4)$$

The Hilbert transform, in practice, acts as a 90º shifter to the signal and the operator generates a real function.

3.1.3 Analytic function ($\breve{s}$)

An analytic signal $\breve{s}\big[n\big]$ is a concept used in continuous time signal processing. It is comprised of a complex time series with real and imaginary parts. Additionally, it has a

Fourier transform that vanishes for negative frequencies (Oppenheim, 1989, Johansson, 2008).

$$\breve{s}[n] = s_{real}[n] + js_{imag}[n]$$ (3.5)

The term analytic can be extended to the *z*-transforms, while the discrete points that they describe in the Argand plane lie inside the unit circle (Oppenheim, 1989) (Leis, 2003).

The analytic signal $\breve{s}(t)$ is defined as the sum of the real part of the signal plus the imaginary part, which is obtained from applying the Hilbert transform $\hat{s}(t)$ to the signal (Kuttruff, 2000).

$$\breve{s}(t) = s(t) + jH[s(t)] = s(t) + j\hat{s}(t)$$ (3.6)

3.1.4 Envelope of a signal ($\overline{s}$)

The use for the Hilbert transform is to help create the envelope of a signal $\overline{s}(t)$ by taking the magnitude of the analytic signal (Kuttruff, 2000).

$$\overline{s}(t) = \sqrt{\left(s(t)\right)^2 + \left(H[s(t)]\right)^2} = \sqrt{\left(s(t)\right)^2 + \left(\hat{s}(t)\right)^2}$$ (3.7)

Applying the envelope to the components of the instantaneous intensity ($\vec{i}_{inst}$) is a good option to locate early reflections, because an envelope of a signal acts like a filter of trivial information without obliterating the vital information (Kuttruff, 2000).

$$\overline{i}_x(t) = \sqrt{\left(i_x(t)\right)^2 + \left(H[i_x(t)]\right)^2} = \sqrt{\left(i_x(t)\right)^2 + \left(\hat{i}_x(t)\right)^2}$$ (3.8)

$$\bar{i}_y(t) = \sqrt{\left(i_y(t)\right)^2 + \left(H\left[i_y(t)\right]\right)^2} = \sqrt{\left(i_y(t)\right)^2 + \left(\hat{i}_y(t)\right)^2} \qquad (3.9)$$

$$\bar{i}_z(t) = \sqrt{\left(i_z(t)\right)^2 + \left(H\left[i_z(t)\right]\right)^2} = \sqrt{\left(i_z(t)\right)^2 + \left(\hat{i}_z(t)\right)^2} \qquad (3.10)$$

The effect on this signal is to smooth the transients of the instantaneous intensity ($\vec{i}_{inst}$) to make it easier to locate the reflections.

The estimation of the angle of the incoming reflected energy $\theta_{i_{xz}}$ may be calculated with simple trigonometry with the envelope of intensity components:

$$\theta_{i_{xz}} = \tan^{-1}\left(\frac{\bar{i}_z(t)}{\bar{i}_x(t)}\right) \qquad (3.11)$$

Therefore, the magnitude of the envelope $\left|\bar{i}_{xz}(t)\right|$ is calculated with the Pythagoras relation:

$$\left|\bar{i}_{xz}(t)\right| = \sqrt{\left(\bar{i}_x(t)\right)^2 + \left(\bar{i}_z(t)\right)^2} \qquad (3.12)$$

In the case of the B-Format signals, Omoto (Omoto and Uchida, 2004) followed the approach of Abdou in (Abdou, 1994). This approach is essentially the same definition of the complex instantaneous intensity defined by Heyser (Heyser, 1986a), but lacking the 1/2 factor that is implying some harmonic averaging. The following expressions generalise the envelope of intensity in orthogonal Cartesian components with the product of the analytic signal taken from the acoustic pressure, multiplied by the conjugate of the associated analytic signal taken from the particle velocity:

$$\bar{i}_x(t) = \left[\breve{p}(t)\right]\left[\breve{\vec{u}}_x(t)\right]^* = \left(p(t) + jH\left[p(t)\right]\right)\left(\vec{u}_x(t) - jH\left[\vec{u}_x(t)\right]\right) \qquad (3.13)$$

$$\vec{i}_y(t) = \left[\breve{p}(t)\right]\left[\breve{\vec{u}}_y(t)\right]^* = \left(p(t) + jH\left[p(t)\right]\right)\left(\vec{u}_y(t) - jH\left[\vec{u}_y(t)\right]\right) \tag{3.14}$$

$$\vec{i}_z(t) = \left[\breve{p}(t)\right]\left[\breve{\vec{u}}_z(t)\right]^* = \left(p(t) + jH\left[p(t)\right]\right)\left(\vec{u}_z(t) - jH\left[\vec{u}_z(t)\right]\right) \tag{3.15}$$

Where:

$\breve{p}$ = analytic acoustic pressure signal.

$\breve{\vec{u}}_x$ = analytic particle velocity vector in *x*-axis.

$\breve{\vec{u}}_y$ = analytic particle velocity vector in *y*-axis.

$\breve{\vec{u}}_z$ = analytic particle velocity vector in *z*-axis.

* = complex conjugate of a complex signal.

The magnitude of the instantaneous intensity ($\vec{i}_{inst}$) is transformed to logarithmic scale for the polar plot by introducing a threshold on the signal at -30 dBFS below the direct sound in the range of 15-30 ms after the direct sound (Begault et al., 2001, Begault et al., 2004, Merimaa and Pulkki, 2005).

3.1.5 Sound intensity ($\vec{i}$)

Sound intensity is an objective acoustic measure that characterises the direction and strength of the rate of flow of energy passing through a unit area per unit time and its measurement unit is {W/m^2}. Sound intensity ($\vec{i}$) is also known as the sound power flux density (Fahy, 1995, Fahy, 2005). In mathematical terms, it is defined as the vector that results after applying the product of the sound pressure (p) and the associated fluid particle velocity vector ($\vec{u}$) (Fahy, 1995).

$$\vec{i}(t) = p(t)\vec{u}(t) \tag{3.16}$$

The measurement of sound intensity ($\vec{i}$) may be done using three different intensity probes principles. a) the p-p probe, b) the *Soundfield* microphone c) the p-u probe (*Microflown* probe). In this thesis, only the face-to-face p-p probe and the *Soundfield* microphone were used. The idea is to capture both the strength and direction of flow as a vector quantity $\left|\vec{i}\right|e^{\theta}$. A single conventional measurement microphone can only measure a scalar quantity called acoustic sound pressure (p) with units of Pascal {Pa}. This non-directional quantity is useful to approximate the sound intensity at a specific position (P) in space, which represents point with coordinates $P(x,y,z)$.

3.1.5.1 Instantaneous intensity ($\vec{i}_{inst}$)

It is an instantaneous vector quantity formed by the product of the instantaneous acoustic pressure and the instantaneous particle velocity vector at any instant on the known factors.

$$\vec{i}_{inst}(t) = \left[p_{inst}(t) \right]\left[\vec{u}_{inst}(t) \right] \qquad\qquad \left\{ \frac{\text{W}}{\text{m}^2} \right\} \qquad\qquad (3.17)$$

For the calculation of the intensity using the p-p intensity probe, some assumptions are used. It is assumed that the intensity is calculated at the geometric centre of the two microphones, which are set in a face-to-face configuration. Therefore, the instantaneous acoustic pressure at the middle point ($\overline{p}_{inst}$) is just the averaged pressure of the two sensors:

$$\overline{p}_{inst}(t) \approx \frac{p_1(t) + p_2(t)}{2} \qquad\qquad \left\{ \frac{\text{N}}{\text{m}^2} = \text{Pa} \right\} \qquad\qquad (3.18)$$

The Euler equation is used to calculate the instantaneous particle velocity ($\vec{u}_{inst}$) in face-to-face p-p probe configuration. It is estimated using a finite difference approximation to the local spatial gradient of sound pressure following expression:

$$\vec{u}_{inst}(t) = -\frac{1}{\rho_0} \int \left(\frac{\partial p(t)}{\partial x} \right) dt \approx -\frac{1}{\rho_0} \int_{-\infty}^{t} \left(\frac{p_1(\tau) - p_2(\tau)}{\Delta x} \right) d\tau \qquad \left\{ \frac{\text{m}}{\text{s}} \right\} \qquad (3.19)$$

While the distance (Δx) is kept constant, it is denoted as the distance (d) between the two pressure sensors so the equation takes the form (Fahy, 1995):

$$\vec{u}_{inst}(t) \approx -\frac{1}{\rho_0 d} \int_{-\infty}^{t} \left[p_1(\tau) - p_2(\tau) \right] d\tau \qquad \left\{ \frac{\text{m}}{\text{s}} \right\} \qquad (3.20)$$

Formerly, the instantaneous intensity component is approximated using the finite difference method by multiplying the estimated instantaneous average pressure ($\bar{p}_{inst}$) by the estimated instantaneous particle velocity ($\vec{u}_{inst}$):

$$\vec{i}_{inst}(t) \approx \frac{1}{2\rho_0 d} \left[p_1(t) + p_2(t) \right] \int_{-\infty}^{t} \left[p_1(\tau) - p_2(\tau) \right] d\tau \qquad \left\{ \frac{\text{W}}{\text{m}^2} \right\} \qquad (3.21)$$

In order to measure the instantaneous direction of flow of sound energy flow ($\vec{i}_{inst}$), at least one pair of phase-matched measurement-grade microphones is needed, as it is defined in the BS EN 61043:1994, IEC 1043:1993 (1994). This may be achieved using more than one pressure sensitive microphone in a probe to measure the sound energy flow ($\vec{i}$) across a single Cartesian axis.

A pair of microphones in a face-to face p-p probe configuration usually measures a couple of acoustic pressures, namely p_1 and p_2. If these two scalar quantities are used to measure the components of the intensity which is co-linear to a single axis, i.e. *x*-axis, the p-p probe needs to be aligned with the line that joins both microphones, and the centre point between the microphones also needs to be aligned with the origin of the Cartesian axis as it is depicted in the left part of **Figure 3.1**.

This creates the first intensity component (i_x). Subsequent measurements of sound intensity along the *y*-axis can be made by rotating the probe 90º from its acoustic centre to

measure acoustic pressures p_3 and p_4. Please refer to **Figure 3.1** to visualise the rotation of the 1-D p-p probe in face-to-face configuration. This procedure creates the second intensity component (i_y). In such a way, it is possible to generate the intensity vector in a plane to create a two-dimensional (2-D) intensity vector ($\vec{i}_{xz}$ or $\vec{i}_{xy}$) i.e. *xz*-plane or *xy*-plane). If a three-dimensional (3-D) intensity vector it is needed, the probe is rotated 90º orthogonal to such plane, to be aligned with the remainder *z*-axis for creating the third component. In that case, acoustic pressures p_5 and p_6 are used to calculate the third intensity component ($\vec{i}_z$) as is shown in **Figure 3.1**.
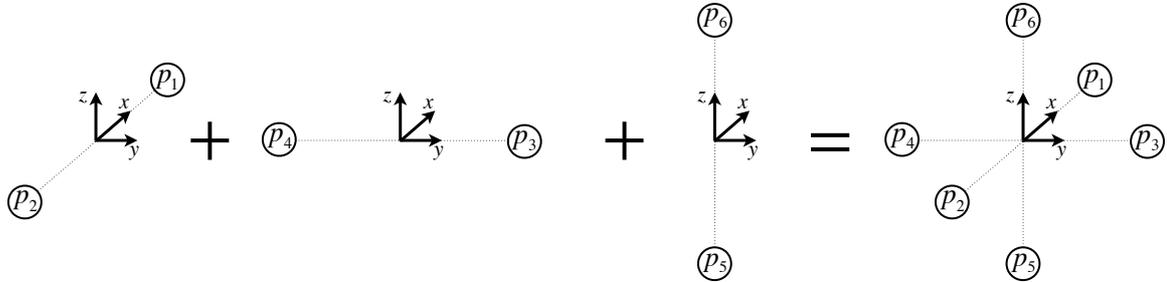
Figure 3.1: Superposition of acoustic intensity ($\vec{i}$) measured with a p-p probe aligned on each Cartesian axis results in 3-D intensity $\vec{i}(x,y,z)$.

The crucial assumption is that the sound field should be time invariant. It basically means that there is not significant change in room temperature and there is no flow of air coming in or out of the room, consequently the general linear equations are valid and superposition of sound fields are summed without any non-linear effects. Further information of theory of sound intensity ($\vec{i}$) implemented on this project is described below.

The following equations are used in the case of calculation instantaneous intensity of a B-Format signal, which is comprised of $w = w(t)$, and $\vec{x} = x(t)\hat{\mathbf{i}}$, $\vec{y} = y(t)\hat{\mathbf{j}}$ and $\vec{z} = z(t)\hat{\mathbf{k}}$, which constitute the 3-D particle velocity vector components:

$$\vec{i}_{inst}(t) = \left[w(t)\right]\left[x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} + z(t)\hat{\mathbf{k}}\right] \qquad \left\{\frac{\text{W}}{\text{m}^2}\right\} \qquad (3.22)$$

3.1.5.2 Active instantaneous intensity ($\vec{i}_a$) in time domain

Active instantaneous intensity ($\vec{i}_a$) is the time-averaged product of the acoustic pressure (p) with the correspondent component of the particle velocity ($\vec{u}$), which has the same phase angle of the acoustic pressure. It represents effective propagation of mean energy flux ($\vec{i}_a$) through space, because it represents the part of the sound intensity, which is not fluctuating in time. It is an indicative figure that shows effective acoustical radiation of sound at a given point in space (Elko, 1984, Fahy, 1995). The direction of the active intensity vector ($\vec{i}_a$) is thus the average direction in which the net energy is flowing and it depends on the relative phase of pressure and particle velocity.

$$\vec{i}_a(t) = \left\langle i_{inst}(t) \right\rangle = \left\langle p(t)\vec{u}(t) \right\rangle = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \left[p(t)\vec{u}(t) \right] dt \qquad (3.23)$$

Reactive intensity ($\vec{i}_{reactive}$) is the time-averaged product of the pressure with the component of the particle velocity, which is 90 degrees apart from the pressure's phase angle. This quantity represents the acoustic power flux ($\vec{i}$) that does not radiate sound and is not transferring sound energy because the phase angle of the pressure ($|p|e^{\theta}$) and the particle velocity $|\vec{u}|e^{\left(\theta + \frac{\pi}{2}\right)}$ are 90 degrees apart. In a phasor diagram, the reactive part is an imaginary term and the active part is the real term and both lie in quadrature (Fahy, 1995).

3.1.5.3 Active instantaneous intensity in frequency domain ($\vec{I}_a$)

In the case of a face-to-face p-p probe embedded in a fixed monochromatic sound fields, the active intensity $\vec{I}_a(\omega)$ is associated with the components of the sound pressure and particle velocity that have the same phase angle as the following relationship (Fahy, 1995, Merimaa et al., 2001, Merimaa and Pulkki, 2005, Merimaa, 2006):

$$\vec{I}_a(\omega) = \mathrm{Re}\left\{\left[P(\omega)\right]^*\left[\vec{U}(\omega)\right]\right\} \qquad \left\{\frac{\mathrm{W}}{\mathrm{m}^2}\right\} \qquad (3.24)$$

Merimaa implemented the case of the calculation of the active instantaneous intensity vector ($\vec{I}_a$) in the B-format context by using a scaling factor of $\sqrt{2}\big/(\rho_0 c_0)$ (Merimaa, 2006). Under the short-time Fourier transform method ($STFT$), each time a window is depicted by two indexes: n denoting time frame and k denoting frequency frame. The active intensity ($\vec{I}_a$) is depicted by:

$$\vec{I}_a(n,k) = \frac{\sqrt{2}}{\rho_0 c_0}\mathrm{Re}\left\{W^*(n,k)\left[X(n,k)\hat{\mathbf{i}} + Y(n,k)\hat{\mathbf{j}} + Z(n,k)\hat{\mathbf{k}}\right]\right\} \qquad (3.25)$$

Where ρ_0 is the air density and c_0 is the speed of the sound, the orthogonal component unit vectors are $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ and $*$ is the complex conjugate operator applied to the complex pressure signal in the B-Format notation (W). The active intensity components are combined to create the three-dimensional active intensity vector:

$$\vec{I}_a(n,k) = \left[I_{a_X}(n,k)\hat{\mathbf{i}} + I_{a_Y}(n,k)\hat{\mathbf{j}} + I_{a_Z}(n,k)\hat{\mathbf{k}}\right] \qquad (3.26)$$

According to Ahonen (Ahonen et al., 2008), the direction of arrival of the active intensity ($\vec{I}_a$) is opposite to the direction of the particle velocity vector with $STFT$ $\vec{U}(n,k)$. The formulas read:

$$\bar{I}_{a_{XY}}(n,k) = \sqrt{\left(-\vec{I}_{a_X}(n,k)\right)^2 + \left(-\vec{I}_{a_Y}(n,k)\right)^2} \qquad \left\{\frac{\mathrm{W}}{\mathrm{m}^2}\right\} \qquad (3.27)$$

$$\theta_{\bar{I}_{a_{XY}}}(n,k) = \left(\frac{180°}{\pi}\right)\tan^{-1}\left(\frac{-\vec{I}_{a_Y}(n,k)}{-\vec{I}_{a_X}(n,k)}\right) \qquad \{°\} \qquad (3.28)$$

$$\phi_{\vec{I}_{a_{XZ}}}\left(n,k\right)=\left(\frac{180^{\circ}}{\pi}\right)\tan^{-1}\left(\frac{I_{a_Z}\left(n,k\right)}{\sqrt{\left(-\vec{I}_{a_X}\left(n,k\right)\right)^2+\left(-\vec{I}_{a_Y}\left(n,k\right)\right)^2}}\right)\qquad\left\{^{\circ}\right\}\qquad(3.29)$$

3.1.5.4 Complex instantaneous intensity ($\vec{i}_{complex}$)

Heyser proposed the time delay spectrometry (TDS) theory in 1967 (Heyser, 1967). It later evolved to include the concept of measurements of dynamic signals by using formulations of complex signals in time domain (TD). The concept of complex instantaneous intensity ($\vec{i}_{complex}$) in time domain was first introduced by Heyser in 1986 in (Heyser, 1986b) and (Heyser, 1986a) as a generalisation of acoustic time delay spectrometry (TDS) theory. It has been supported by Jacobsen (Jacobsen, 1991), Fahy (Fahy, 1995) and Torres-Guijarro (Torres-Guijarro et al., 2011) as a good technique to analyse transient complex signals.

Further analysis for the report involves the use of complex instantaneous intensity calculation according to Heyser's formula (Heyser, 1986a):

$$\vec{i}_{complex}\left(t\right)=\frac{1}{2}\left[p_{complex}\left(t\right)\right]\left[\vec{u}_{complex}\left(t\right)\right]^{*}=\frac{1}{2}\left(p\left(t\right)+jH\left[p\left(t\right)\right]\right)\left(\vec{u}\left(t\right)-jH\left[\vec{u}\left(t\right)\right]\right)\quad(3.30)$$

Where:

$\vec{i}_{complex}$ = complex instantaneous intensity in time domain.

* = complex conjugate operator.

$p\left(t\right)$ = acoustic pressure.

$u\left(t\right)$ = particle velocity vector.

$H\left[\vec{u}\left(t\right)\right]$ = Hilbert transform of the particle velocity vector.

$H\big[\,p(t)\big]$ = Hilbert transform of the acoustic pressure.

The analytic signal of the pressure ($\breve{p}$) and the particle velocity ($\breve{\vec{u}}$) are formed by using the sum of the real part of the signal plus the imaginary part, which is formed by multiplying the Hilbert transform of the signal by the complex operator (j).

$$\breve{\vec{u}}(t) = \vec{u}(t) + jH\big[\,\vec{u}(t)\big] \tag{3.31}$$

$$\breve{p}(t) = p(t) + jH\big[\,p(t)\big] \tag{3.32}$$

Active complex instantaneous intensity ($\vec{i}_{\substack{active\\complex}}$) is defined as the real part of the complex instantaneous intensity :

$$
\begin{aligned}
\vec{i}_{\substack{active\\complex}}(t) &= \mathrm{Re}\left\{\frac{1}{2}\big[\,p_{complex}(t)\big]\big[\,\vec{u}_{complex}(t)\big]^{*}\right\}\\[4pt]
&= \mathrm{Re}\left\{\frac{1}{2}\big[\,p(t) + j\hat{p}(t)\big]\big[\,\vec{u}(t) - j\hat{\vec{u}}(t)\big]\right\}\\[4pt]
&= \mathrm{Re}\left\{\frac{1}{2}\big(p(t) + jH\big[\,p(t)\big]\big)\big(\vec{u}(t) - jH\big[\,\vec{u}(t)\big]\big)\right\}
\end{aligned} \tag{3.33}
$$

3.1.6 Diffuseness estimate (ψ)

Merimaa, (Merimaa, 2006) proposed a measure called diffuseness estimate (ψ), which was used to characterise the decay of sound as a ratio of the nearby restricted energy. It was based on the proposed field indicator by Stanzial et al. called pressure velocity correlation coefficient (η) (Stanzial et al., 1996, Stanzial et al., 2000), based on active instantaneous intensity ($\vec{i}_{a}$) (Stanzial and Prodi, 1997). This measure is used to compare performance of intensity probes reported as the radiation index (η) by Cengarle (Cengarle, 2012).

$$\eta = \frac{\left\| \left\langle \vec{I}_a \right\rangle \right\|}{\left\langle cE \right\rangle} \qquad \qquad \left\{ \text{adimensional} \right\} \tag{3.34}$$

The degree of diffuseness of a sound field is a scalar value, which measures the ratio between arriving directions of active intensity ($\vec{I}_a$) and acoustic energy density (E) at a given time window. Pulkki (Pulkki, 2007) explains that diffuseness varies between values of 0, when the direction of the incoming reflections directions is clearly determined i.e. an isolated reflection and 1, when the direction of the incoming reflections is random (implying a diffuse field). In **Figure 6.57**, the darker grey spot values correspond to higher values and the white values to lower values of diffuseness estimate (ψ).

The Potential acoustic energy in frequency domain (E_{pot}) is calculated using the following expression:

$$E_{pot}(\omega) = \frac{\left[P(\omega) \right]^2}{2\rho_0 c_0^{\ 2}} \qquad \qquad \left\{ \frac{\text{kg}}{\text{m s}^2} = \text{Pa} \right\} \tag{3.35}$$

The kinetic acoustic energy in frequency domain (E_{kin}) can be formulated as:

$$E_{kin}(\omega) = \frac{1}{2}\rho_0 \left| \vec{U}(\omega) \right|^2 \qquad \qquad \left\{ \frac{\text{kg}}{\text{m s}^2} = \text{Pa} \right\} \tag{3.36}$$

The acoustic energy density (E) is calculated with the sum of potential energy (E_{pot}) and acoustic kinetic energy (E_{kin}), which in the frequency domain can be written as:

$$E(\omega) = \frac{1}{2}\left(\frac{\left[P(\omega) \right]^2}{\rho_0 c_0^{\ 2}} + \rho_0 \left| \vec{U}(\omega) \right|^2 \right) \qquad \left\{ \frac{\text{kg}}{\text{m s}^2} = \text{Pa} \right\} \tag{3.37}$$

According to Merimaa (Merimaa, 2006) the diffuseness estimate (ψ) in a condensed form is obtained by applying the Euclidean norm ($\| \ \|$) to the averaged active instantaneous intensity using the *STFT* ($\vec{I}_a$), and then divided by the averaged product of the energy density (E) multiplied by the speed of sound (c). This quotient is bonded between 0 (non-diffuse) and 1 (diffuse) by taking its complement. This inverted scale is performed by subtraction of the mentioned quotient from a unity value.

$$\psi = 1 - \frac{\left\| \left\langle \vec{I}_a \right\rangle \right\|}{\left\langle cE \right\rangle} \qquad \left\{ \text{adimensional} \right\} \qquad (3.38)$$

Ahonen (Ahonen and Pulkki, 2009) recently proposed a new expression to evaluate diffuseness estimate ($\hat{\psi}$) by using only the instantaneous active intensity ($\vec{i}_a$) in the time domain. A comparison of diffuseness estimates was performed by De Galdo and Ahonen (Del Galdo et al., 2012), who recognise the need to perform spatial averaging to ameliorate the variance on these estimates. However, both estimates show similar performance. Before Merimaa (Merimaa, 2006), it was defined in terms of a quotient of the length of the active instantaneous intensity in frequency domain ($\vec{I}_a$) upon the energy density (E). The process to calculate the proposed diffuseness estimate ($\hat{\psi}$) is dividing the length of averaged active instantaneous intensity vectors $\left\| \left\langle \vec{i}_a(t) \right\rangle \right\|$ by the averaged length of the active instantaneous intensity $\left\langle \left\| \vec{i}_a(t) \right\| \right\rangle$ and then taking its complement and performing the square root. It is depicted in the following expression:

$$\hat{\psi}(t) = \sqrt{1 - \frac{\left\| \left\langle \vec{i}_a(t) \right\rangle \right\|}{\left\langle \left\| \vec{i}_a(t) \right\| \right\rangle}} \qquad \left\{ \text{adimensional} \right\} \qquad (3.39)$$

The new proposed temporal-variation based diffuseness estimate ($\hat{\psi}$) apparently produces more reliable diffuseness estimation than the traditional method (ψ) in one and two-dimensions, so it may help to locate reflections with more accuracy. Pihlajamäki

tested the new formulation for diffuseness estimate denoted as ($\hat{\psi}$) in (Pihlajamäki, 2009).

In order to apply the diffuseness estimate (ψ) to the case of B-Format signals, the following approach is taken:

The acoustic energy density (E) in the spectrum applied for B-Format signal has the following form:

$$E(\omega) = \frac{1}{2}\frac{1}{\rho_0 c^2}\left[|W(\omega)|^2 + \frac{1}{2}|X(\omega)\hat{\mathbf{i}} + Y(\omega)\hat{\mathbf{j}} + Z(\omega)\hat{\mathbf{k}}|^2\right]$$ (3.40)

Further development from applying the analysis to B-Format signals led to an instantaneous function of diffuseness (ψ_{inst}) in terms of short-time Fourier transform windows ($STFT$) in a room by applying the following expression (Enroth, 2007), (Merimaa, 2006):

$$\psi_{inst}(n,k) = 1 - \left[\frac{\sqrt{2}\left\|\text{Re}\left\{W^*(n,k)\vec{U}(n,k)\right\}\right\|}{|W(n,k)|^2 + \frac{1}{2}\left(\left|-\vec{I}_{a_X}(n,k)\right|^2 + \left|-\vec{I}_{a_Y}(n,k)\right|^2 + \left|-\vec{I}_{a_Z}(n,k)\right|^2\right)}\right]$$ (3.41)

The instantaneous value of diffuseness (ψ_{inst}) is calculated by applying the concept of the 'Directional Diffusion Index' (ψ) to the active instantaneous intensity in frequency domain ($\vec{I}_a$) orthogonal component vectors calculated from B-format signal.

The *Matlab*'s '`norm`' operation calculates the Euclidean length (l^2 norm) of a 3-D vector and is applied to the active intensity vector in the frequency domain ($\vec{I}_a$) in the following way:

$$\left\|\vec{I}_a(n,k)\right\| = \sqrt{\left(\vec{I}_{a_X}\right)^2 + \left(\vec{I}_{a_Y}\right)^2 + \left(\vec{I}_{a_Z}\right)^2}$$ (3.42)

To the 'expected value' of intensity $\left\langle \vec{I}_a(n,k) \right\rangle$, which in practice is calculated with the recursive integration of the short-time Fourier transform ($STFT$) of the active instantaneous intensity vector $\vec{I}_a(n,k)$.

3.1.7 Equivalent rectangular bands (ERB)

Anatomically speaking, the equivalent rectangular bands represents sections of length approximately of 1.3 mm, which imaginarily divide inside our basilar membrane. The equivalent rectangular bands act as a filter bank consisting of a series of band pass filters, which are frequency dependent. Inside a critical band our hearing system responses to frequency changes immediately and it is the responsible of masking sounds, which are outside each equivalent rectangular band (Vaseghi, 2007). This filter banks are approximated by a rectangular shape, however in reality an accurate model is curved (Slaney, 1993). In order to obtain the directions of the reflections, a short-time Fourier transform ($STFT$) analysis is made to the intensity vector in the horizontal plane. The time window duration is chosen according to the sampling frequency and the mean free path (l_{MFP}) of the room. Octave band, 1/3 octave band, or equivalent rectangular band (ERB) filtering can be used to smooth the time windows. The best perceptual approximation is the ERB because it resembles human hearing resolution, whereas 1/3 octave band approximates the critical band with less accuracy and octave band may be useful for engineering analysis. The formula used to calculate the bandwidth of the rectangular filters (ERB_N) is taken from Moore's latest formulation (Moore, 2003):

$$ERB_N = 24.7\left(4.37 \times 10^{-3} f + 1\right) \qquad \left\{ \mathrm{Hz} \right\} \qquad (3.43)$$

The variation of the critical band's bandwidth can be treated as a function of frequency (Vaseghi, 2007):

$$BW_c(f) = 25 + 75\left[1 + 1.4\left(f \times 10^{-3}\right)^2\right]^{0.69} \qquad \left\{ \mathrm{Hz} \right\} \qquad (3.44)$$

3.2 Statistical treatment for the sample of angular measurements

3.2.1 Uncertainties

According to the "GUM: Guide to the Expression of Uncertainty in Measurement", the definition of 'uncertainty' is related to the doubt about the validity of the result of the measurement (2008). The definition of the 'uncertainty of measurement' is related to the result of a statistical measurement, which reports the dispersion of the values that could reasonably be attributed to the measured parameter. It is usually reported as a standard deviation. The confidence interval needs to be constructed according to the size of the sample. For small samples below 30 measures, the normal distribution may be approximated by the Student´s t-distribution (Navidi, 2006).

Time invariance is a topic that can be measured with the standard deviation of a sample of measurements performed on the same position.

The detection of the angle of arrival of direct sound (θ_{s_i}) is trivial, but the detection of the angle of arrival of first reflection (θ_{m_i}) is not so, and usually exhibits larger errors in its estimation. Therefore, it is expected that the first reflection will exhibit larger error and variation of the positions, particularly as the reflection paths become longer. Since the high sensitivity of the system to reflection angle and its natural tendency to increase errors in estimation, two measurement methods were used; the short-time Fourier transform ($STFT$) and the time domain methods (TD).

3.2.2 Linear Error (ε) and linear standard deviation(σ)

Most of the graphs that were planned to be results in the report were done with simple error (ε) and its dispersion statistics (σ_ε). It was applied as linear statistics and circular statistics depending on the case studied. The input variables used were: a) spatial position

varied in a single dimension (*x*- axis) and b) time variance measured as the standard deviation of the error (σ_ε) found in the estimation and the ith measured angles of a sample (θ_{m_i} and θ_{s_i}). In the case of the p-p intensity probe has become a problem to take into account when three different rotations of the probe are performed.

According to Tashev (Tashev, 2009), the evaluation of sound localisers is done using the following parameters: mean error ($\bar{\varepsilon}$) and the standard deviation error also known as mean squared value (σ_ε):

$$\bar{\varepsilon} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\varphi}_i - \varphi_i) \qquad \left\{ {}^\circ \right\} \qquad\qquad (3.45)$$

The overall linear mean error can be computed by performing an arithmetic mean of the error ($\bar{\varepsilon}_{\bar{\varepsilon}}$) taken from the ith error samples, and takes the form:

$$\bar{\varepsilon}_{\bar{\varepsilon}} = \frac{1}{N}\sum_{i=1}^{N}\bar{\varepsilon}_i \qquad \left\{ {}^\circ \right\} \qquad\qquad (3.46)$$

The unbiased standard deviation error of a sample is calculated when the sample size is below 30 measurements by using the following expression:

$$\sigma_{\bar{\varepsilon}} = \sqrt{\left(\frac{1}{N-1}\right)\sum_{i=1}^{N}(\hat{\varphi}_i - \varphi_i)^2} \qquad \left\{ {}^\circ \right\} \qquad\qquad (3.47)$$

Where:

$\bar{\varepsilon}_{\bar{\varepsilon}}$ = overall mean error. $\left\{ {}^\circ \right\}$

$\bar{\varepsilon}$ = mean error $\left\{ {}^\circ \right\}$.

N = number of samples $\{\ \ \}$.

$\hat{\varphi}_i$ = sound source i^{th} estimated angle $\left\{ {}^\circ \right\}$.

φ_i = sound source i^{th} real angle $\{\,^{\circ}\,\}$.

The application of the correction factor under the Student´s t-distribution uses the following formula (Abramowitz and Stegun, 1972):

$$\sigma_{\varepsilon_{corrected}} = t_{N-1,\alpha/2}\left(\frac{\sigma_\varepsilon}{\sqrt{N}}\right) \tag{3.48}$$

Where:

σ_ε = standard deviation of a unbiased sample of errors

N = sample size or degrees of freedom

$t_{N-1,\alpha/2}$ = percentage points of Student's t-distribution

For a $N = 10$, if a confidence interval with 0.95 probability of being inside the probability distribution is chosen, then the value of $t_{N-1,\alpha/2} = 1.833$.

3.2.3 Linear Percentage error ($\varepsilon_{\%}$)

The definition of linear percentage error is the quotient of the error (ε) divided by the real value (φ) expressed in percentage.

According to Abramowitz and Stegun (Abramowitz and Stegun, 1972), the formula for percentage error adapted for this report is:

$$\varepsilon_{\%} = \left(\frac{\Delta x}{x}\right)\times 100 = \left(\frac{\varepsilon}{\varphi}\right)\times 100 \qquad \{\%\} \tag{3.49}$$

Where:

$\varepsilon = \hat{\varphi} - \varphi$ = error. $\{\,^{\circ}\,\}$

φ = real angle. $\{\,^{\circ}\,\}$

$\varepsilon_{\%}$ = percentage error. {%}

3.2.3.1 Absolute percentage error $\left| \varepsilon_{\%} \right|$

This parameter is used more treating linear errors, considering that errors need to be positive in order to compare them. It is simply the absolute value of the percentage error ($\varepsilon_{\%}$):

$$\left| \varepsilon_{\%} \right| = \left| \left(\frac{\varepsilon}{\varphi} \right) \times 100 \right| \qquad \left\{ \% \right\} \qquad (3.50)$$

3.2.3.2 Mean Absolute Percentage Error $\left| \bar{\varepsilon}_{\%} \right|$ or ($MAPE$)

The mean absolute percentage error ($MAPE$) $\left| \bar{\varepsilon}_{\%} \right|$ also known as mean absolute deviation is used to analyse the behaviour of time series in financial reports as a forecast error (Hamilton, 1994). For the generation of graphs in this report (see chapter 6), it is implemented by taken the linear mean of the ith samples of absolute percentage error $\left| \varepsilon_{\%} \right|$, and it takes the form:

$$\left| \bar{\varepsilon}_{\%} \right| = \frac{1}{N} \sum_{i=1}^{n} \left| \varepsilon_{\%_i} \right| \qquad \left\{ \% \right\} \qquad (3.51)$$

3.2.3.3 Weighted average of Standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$)

The weighted average of standard deviation error of $MAPE$ ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) is chosen when there are different uncertainties on each sample. It may be treated as having different statistical variables (Navidi, 2006). The weighted average is calculated taken the sum of the ith samples of variance of the error ($\sigma_{\bar{\varepsilon}_{\%}}^{2}$) and then taking the square root of it and divide it by the number of samples (N):

The correspondent weighted standard deviation used for describing the uncertainty of a sample of measurements is adapted to use the mean absolute percentage error $\left|\bar{\varepsilon}_{\%}\right|$ using the following expression:

$$\bar{\sigma}_{\bar{\varepsilon}_{\%}} = \frac{1}{N} \sqrt{\sum_{i=1}^{N} \left(\sigma_{\bar{\varepsilon}_{\%_i}} \right)^2} \qquad \{\%\} \tag{3.52}$$

This expression is a good option for rejection of any outliers in the measurements, because it reports a smaller value than the traditional mean of standard deviation ($\bar{\sigma}$). In this particular case, the variance (σ^2) is arithmetically averaged, then is converted to standard deviation (σ) by taking the square root of the average of variances.

3.2.3.4 Overall Mean Absolute Percentage Error ($OMAPE$) $\left|\bar{\varepsilon}_{\bar{\varepsilon}_{\%}}\right|$

The overall mean absolute percentage error can be computed using linear statistics by performing an arithmetic mean of the error $\left|\bar{\varepsilon}_{\bar{\varepsilon}_{\%}}\right|$ taken from the ith error samples, and it takes the following expression:

$$\left|\bar{\varepsilon}_{\bar{\varepsilon}_{\%}}\right| = \left| \frac{1}{N} \sum_{i=1}^{N} \bar{\varepsilon}_{\%_i} \right| \qquad \{\%\} \tag{3.53}$$

3.2.4 Directional statistic parameters

In order to calculate the mean direction over a measured angle, the angle data in degrees need to be converted to radians prior to any calculation because the *Matlab* trigonometric functions calculate data in radians by multiplying by the factor $\left(\pi \text{ rad}\right)/180°$. Let the sample of angles be $\theta_1, \theta_2, \theta_3, ..., \theta_N$. The resultant direction or mean direction ($\bar{\theta}$) is calculated using the following formula proposed in (Trauth, 2010). The values of cosine and sine are decomposed for each direction θ_i to analyse it by its resultant sum of

Cartesian components (x_r, y_r) and are then used to calculate the resultant or mean direction ($\bar{\theta}$) for the set of angular data.

$$x_r = \sum_{i=1}^{N} \cos\theta_i \qquad \{\text{adimensional}\} \qquad (3.54)$$

$$y_r = \sum_{i=1}^{N} \sin\theta_i \qquad \{\text{adimensional}\} \qquad (3.55)$$

3.2.4.1 Circular mean ($\bar{\theta}$)

The resultant direction ($\bar{\theta}$) of a sample is called circular mean and is calculated with the trigonometric relationship:

$$\bar{\theta} = \tan^{-1}\left(\frac{y_r}{x_r}\right) \qquad \{\text{rad}\} \qquad (3.56)$$

The length of the resultant (R) is computed with the Pythagoras relation applied to the

$$R = \sqrt{(x_r)^2 + (y_r)^2} \qquad \{\text{adimensional}\} \qquad (3.57)$$

The length of the resultant (R) evidently varies with the dispersion of the sample. Applying the normalisation to the resultant length to the number of measurements produces the mean resultant length ($\bar{R}$):

$$\bar{R} = \frac{R}{N} \qquad \{\text{adimensional}\} \qquad (3.58)$$

3.2.4.2 Circular variance (σ_o^2)

The circular variance (σ_o^2) is calculated with the following expression:

$$\sigma_o^{\,2} = 1 - \bar{R} \qquad \{rad\} \qquad (3.59)$$

3.2.4.3 Circular standard deviation (σ_o)

The circular standard deviation is calculated by two different expressions the angular deviation (Berens, 2009):

$$\sigma = \sqrt{2\bar{R}} \qquad \{rad\} \qquad (3.60)$$

Which is bounded between 0 and $\sqrt{2}$ and the Mardia's circular standard deviation formula (Fisher, 1971):

$$\sigma_o = \sqrt{-2\log(\bar{R})} \qquad \{rad\} \qquad (3.61)$$

Which is bounded between 0 and ∞. In order to convert the circular data to degrees it is multiplied by the factor $180°\big/(\pi\;rad)$. Both expressions deliver similar results, where the circular deviation tends to report slightly smaller values, therefore in this thesis the circular deviation is used (σ_o) when angles of arrival (θ) are analysed.

3.3 Summary

This chapter covers the applied intensimetry theory and the calculation methods for the two probes chosen, (i.e. face-to-face p-p probe and *Soundfield* microphone). It also covers the statistical treatment for the measurements of angles. The first topic explains the concept of the four-channel orthogonal impulse responses $\mathbf{P}_4(\vec{r},t)$. In order to explain the basics to obtain the advanced intensity quantities, explanation of the following concepts are covered: the Hilbert transform $H\big[s(t)\big]$, the analytic signal ($\breve{s}$), and the envelope of the signal $\bar{s}(t)$ (Heyser, 1986a, Kuttruff, 2000, Johansson, 2008). The properties of the envelope help to locate early reflections with more accuracy.

Consequently, it is applied to the three-dimensional impulse response and direction of arrival (DOA) and time of arrival (TOA) and strength is computed from it. Subsequently, the sound intensity theory is introduced. It comprises of explanations of instantaneous intensity ($\vec{i}_{inst}$), active instantaneous intensity ($\vec{i}_a$), and Heyser's complex intensity ($\vec{i}_{complex}$) in time domain, which tends to improve the estimation of the reflections because of its properties of averaging the intensity. The useful part of the complex part is called active complex instantaneous intensity ($\vec{i}_{active_{complex}}$) and is the part used in the extraction of directional information. There are different approaches taken to calculate this quantities depending on the type of probe chosen and the domain chosen. The following topic to cover is the diffuseness estimate (ψ) and finally the implementation of equivalent rectangular bands, which are used to resemble our spatial resolution of the hearing system. The next section covered is the statistical treatment for the sample of angular measurement, which is comprised of uncertainties, linear errors (ε) and standard deviation (σ), linear percentage error ($\varepsilon_{\%}$) and directional statistic parameters for angular quantities, which consist of circular mean ($\bar{\theta}$), dispersion of the sample ($\bar{R}$), circular variance ($\sigma_o{}^2$) and circular standard deviation (σ_o).

Chapter 4 : Acoustic Measurement System

The detection of direct sound has not been a problem for microphone arrays. It is usually comprised of the simple task of finding where it shows its maximum level, with the exception of cases where there are strong echoes that can be detected in the measuring position in the room. In these cases, it is possible to encounter an acoustically amplified reflection, which can exhibit a higher level than the level of the direct sound, because of the phenomena of focusing sound reflections, which may happen when sound encounter a concave surface (Ballou, 1991). Depending on the size of the room, it usually has a certain time gap from early reflections. Generally, the only concern for locating the direct sound is to be sure that the measurement system is sample accurate in all the channels, and that the microphone capsules are calibrated to have the same sensitivity. This sound localization task can be done either using the particle velocity estimates ($\vec{u}$) from the raw data signals, or by calculating the sound intensity ($\vec{i}$) from them. The latter method is recommended, since sound intensity ($\vec{i}$) is proportional to the acoustic energy (E), and can be used for the following analysis of early reflections without bias. The particle velocity estimates have a disadvantage of not being able to show where the real flow of sound energy is. This happens especially in reverberant environments (Jacobsen, 1979), where there are strong and fast changes of sound energy flow ($\vec{i}$) (Jacobsen, 2002), or in the case of measuring in non-diffuse spaces, where the pressure fluctuates in a dissimilar manner with respect to the sound intensity ($\vec{i}$). The reason for this behaviour is because both signals are uncorrelated (Gade, 1985). Despite this argument, care needs to be taken with the difference between uncorrelated and incoherent signals discussed by Jacobsen in (Jacobsen and Nielsen, 1987).

The contributions of the author are in the design of the custom-made rotation cradle, and the post-processing and analysis of the measured data in order to use different types of sound intensities to characterise the strength, spatial and temporal attributes of the early reflections in small rooms.

4.1 Implementing the physical measurement system

The core of this project was originally established to develop a measurement system aiming to measure the strength, temporal and spatial properties of early sound fields in small rooms with a spherical resolution similar to the minimum audible angle (MAA = ± 2º at the front (Mills, 1958, Yost, 1974)).

It is also called localization blur (Blauert, 1997a). Our ability to perceive spatial differences is not constant across the horizontal and vertical plane (see **Figure 4.1**). For the horizontal plane, the minimum audible angle can be determined at the front with a maximum accuracy of 1 degree using white noise signals. Localization blur fluctuates with different signals stimuli.
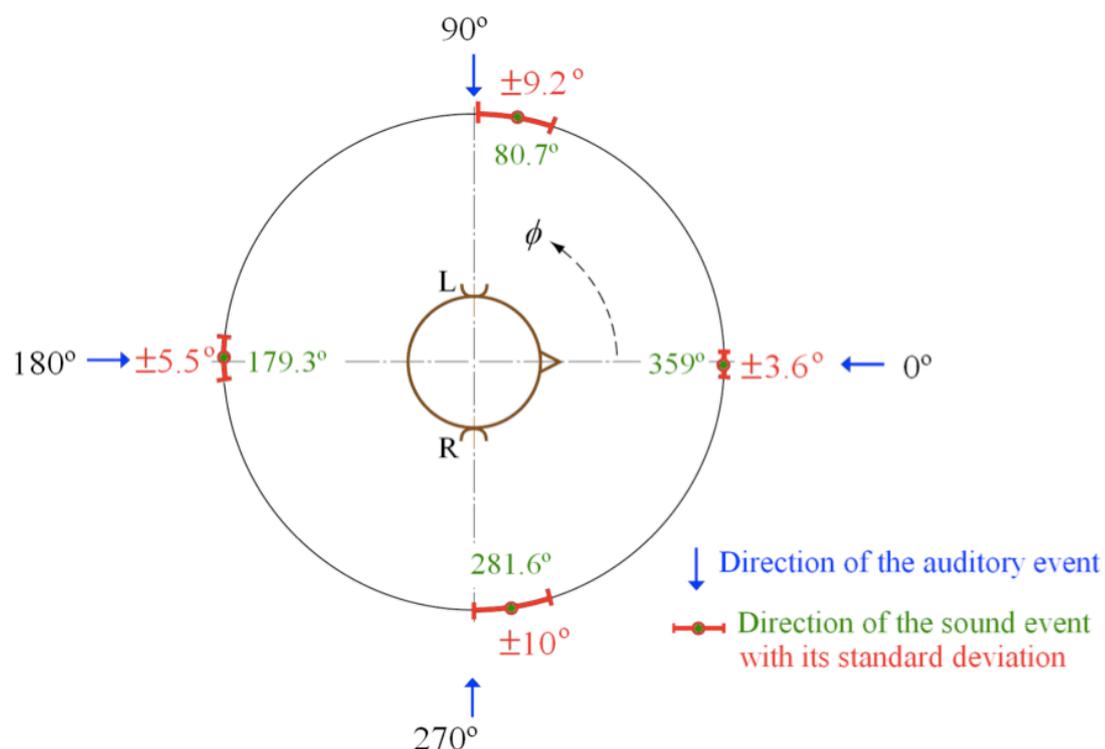
Figure 4.1: Localisation blur in horizontal plane (after Blauert)

These objectives are fulfilled by carrying out a number of experiments, which consist of obtaining a three-dimensional impulse response (3-D IR) of a reflection's controlled environment. The results of these measurements are post-processed to obtain energetic

acoustic quantities based in sound intensity ($\vec{i}$). The results are compared with the known reference, which comprises geometrical information in order to perform accuracy and error analysis. This is achieved by using carefully chosen advanced metrology techniques. A comparison of a couple of acoustic probes was done against the geometrical reference. Later, the post-processing data was also compared with two sound intensity ($\vec{i}$) analysis techniques based in time domain and frequency domain. The probes used were the p-p intensity probe and the *Soundfield* microphone.

The spatial measurement method is based on the use of precision optical solutions based on a couple of orthogonal positioned laser cross devices. The physical alignment of the system is possible using this configuration and its function is to enhance the repeatability of the proposed method. Initially, the accuracy of the positions measured was done using linear rulers and plumbs. It soon, became necessary to improve the accuracy of the system by adding wheels to the triplay base of the mobile tripod that holds the acoustic probe. Nevertheless, the achieved accuracy using these solutions was not state-of-the- art and it was planned to implement a system with higher repeatability and accuracy based on better equipment.

Previous experience of the author working on the automation of a tube cutter machine and on the implementation of a vertical machine center (*VMC*) using servo-motors, provides the insight to find an appropriate solution for the accurate positioning. The use of linear guide ways borrowed from the *CNC* machinery components[4] can help to move the receiver with accuracy and repeatability for consistent positioning. If additional accuracy is needed, the measuring method can use a servomotor for linear location of the source with the highest repeatability down to ±1 μm. Unfortunately, this idea could not be implemented in the measurements performed for this report because of the unavailability of the linear guides during the measurements.

The post-processing of the data acquired needs to be the key part of the innovation of the method's aims. The measurement method has tried to select the most economical choices in terms of number of channels by using new quantities of sound intensity borrowed from

---

[4] https://tech.thk.com/en/products/pdfs/en_a01_350.pdf [Online accessed on May 19th 2013]

Heyser's time delay spectrometry (TDS) research (Heyser, 1967), where he first introduces the Energy Time Curve (ETC). Later, he proposes the use of complex instantaneous intensity ($\vec{i}_{complex}$) (Heyser, 1986a). A higher resolution in acquisition of data has been addressed in order to obtain the most detailed transient response in the time domain, which later can offer detail in the frequency domain analysis. The aim of this proposed measurement method has borrowed the multidimensional display of frequency and directional information of early reflections in order to be able to find a good amount of orthogonal variables to present the analysis of the sound field as a 2-D plane description of the sound decay of a given room. This method can be used in large rooms and small rooms by adjusting the time windows. The portability of the equipment should be an important issue for the practitioners that may be interested to assess the 3-D transient behaviour of rooms, because a practical measurement system needs to be easy to use and manageable. In the future, real-time analysis of early-reflected energy may be commercially implemented in products music such as the LMS *SoundBrush*[5] or *Eigenmike*[6] (Meyer and Elko, 2010). It is particularly difficult to address the accuracy of the method for measuring the low-frequency content of the reflections that characterise small rooms, since the modal activity that distorts the sound field and creates a non-diffuse sound field. Nevertheless, the method needs to take into account limitations on low frequency and a combination of the more accurate results that can be achieved when mid and high frequencies are measured. However, the p-p type sound intensity probe has a limited range of frequency applicability when using a single, solid, plastic size-spacer minimising shadow, and reflection effects. Changing and aligning the acoustic centre is crucial to maintain directional accuracy and is time consuming. It inherently limits the measurement system proposed by the compromise of using the adequate dimensions spacers. The *Microflown* intensity probe can overcome these limitations because it is broadband, nevertheless has a limited dynamic range at high frequencies (Cengarle, 2012), which is not as high as the p-p intensity probe's dynamic range (Jacobsen and de Bree, 2005).

---

[5] http://www.lmssoundbrush.com/discover and http://www.lmssoundbrush.com/product/antenna
[Online accessed on March 19th 2013]
[6] http://www.mhacoustics.com/mh_acoustics/Eigenmike_microphone_array.html
[Online accessed May 23rd 2013]

4.2 Early reflection real values calculation method

The initial aim was to measure five positions beginning with a calibration position, which was set at an arrival angle of direct sound at 45º. This value was convenient because for Soundfield and p-p probe has symmetric components in both *x* and *y*-axis. It was assumed that the sound travels as a ray. This assumption is valid only for wavelengths that are smaller than the surface where they are reflected. In practice this happens just for mid and high frequencies.

4.2.1 Input data and output data calculation

The ith angle of arrival of first reflection (θ_{m_i}) is denoted with the sub-index *m,* which means arriving directly at the microphone   is selected arbitrarily by the angular step used in section 5.2.1 and its calculation is made with elemental trigonometry applied to the rectangular triangles presented in **Figure 5.8**

$$\theta_{m_i} = \tan^{-1}\left(\frac{h_m + h_s}{x_i}\right) \qquad \left\{^\circ\right\} \tag{4.1}$$

where:

θ_{m_i} = angle of arrival of first reflection at the mic $\left\{^\circ\right\}$

h_m = height of the microphone $\left\{\text{m}\right\}$

h_s = height of the source (studio monitor) $\left\{\text{m}\right\}$

x_1 = distance on the x axis $\left\{\text{m}\right\}$

In order to calculate the ith distances (x_i) in the *x*-axis which create the desired angle of arrival first reflection at a given position (P_i), the following equation was used:

$$x_i = \frac{\left(h_m + h_s\right)}{\tan\left(\theta_{m_i}\right)} \qquad \left\{\text{m}\right\} \tag{4.2}$$

Where:

x_i = microphone position in x axis $\{m\}$

h_m = microphone height $\{m\}$

h_s = source height $\{m\}$

θ_{m_i} = angle of the first reflection if the ith position $\{°\}$

The increments of distance along *x*-axis for the generation of the step angles are calculated by:

$$\Delta x_i = |x_{i-1} - x_i| \qquad \{m\} \qquad\qquad (4.3)$$

These distance calculations have the maximum length for the set of measurements, which is possible within the dimensions of the semi-anechoic chamber of the Acoustic laboratory of the University of Salford. These inner dimensions are $4.170 \text{ m} \times 3.270 \text{ m} \times 2.950 \text{ m}$. It covered a range of angles (θ_{s_i}) from 45.000º down to 14.966º (please refer to **Figure 5.10**). The computation of the ith angles of arrival of the direct sound (θ_{s_i}), where the sub-index *s* means arriving from first reflection from the source, use the *x*-axis coordinate values by using the following formula:

$$\theta_{s_i} = \tan^{-1}\left(\frac{h_s - h_m}{x_i} \right) \qquad \{°\} \qquad\qquad (4.4)$$

Where:

θ_{s_i} = angle of the direct sound (source)

The distance travelled of the direct sound from the source to the receiver is calculated with the rectangle triangle formula:

$$r_{i_{direct}} = \sqrt{\left(h_s - h_m\right)^2 + \left(x_i\right)^2} \qquad \{m\} \qquad\qquad (4.5)$$

Where:

$r_{i_{direct}}$ = distance of the direct sound in position i^{th} $\{m\}$

h_s = height of the source $\{m\}$

h_m = height of the source $\{m\}$

x_i = distance from x axis from the source to the receiver $\{m\}$

Now the time delay of the ith direct sound is calculated by the following formula:

$$t_i = \frac{r_{i_{direct}}}{c_0} = \frac{\sqrt{(x_i)^2 + (h_s - h_m)^2}}{c_0} \qquad \{s\}$$

(4.6)

4.2.2 Time of flight (TOF)

The time difference (also known as time of flight or (TOF)) from the first reflection to the direct sound is calculated as the measure to assess the reliability of the impulse responses.

$$\Delta t_{i_{dir-ref}} = t_{i_{reflection}} - t_{i_{direct}} \qquad \{s\}$$

(4.7)

Where:

Δt_i = Time Delay of i_{th} position of first reflection $\{s\}$

$t_{i_{reflection}}$ = Time of flight of the first reflection $\{s\}$

$t_{i_{direct}}$ = Time of flight of the direct sound $\{s\}$

Expanding the expression in terms of distances yield:

$$\Delta t_{i-1} = \frac{\sqrt{(x_i)^2 + (h_m + h_s)^2} - \sqrt{(x_1)^2 + (h_s - h_m)^2}}{c_0} \qquad \{s\}$$

(4.8)

It is important to note that the positions that are far from the source tend to have smaller differences of time, and its impulse response (*IR*) will show that compressed version between the direct sound and the first reflection.

Calculating the difference of adjacent position's time delay is possible to provide a measure for proper selection of the desired minimum time resolution. It is well suited for measuring the changes of position:

$$\Delta t_{i \rightarrow i+n} = \left| \Delta t_{i+n} - \Delta t_i \right| \qquad \qquad \left\{ s \right\} \qquad \qquad (4.9)$$

4.2.3 Exponential sine sweep method

The impulse response of the room represents the interaction of the speaker and the room at this particular position. It is important to note that these impulse responses are not good to calculate any acoustic parameter, because studio monitors are not omni-directional at all frequencies, as even a purpose-designed, omni-directional source is limited to mid frequencies (2012)[7]. The stimulus signal chosen is logarithmic swept sine, given that the robustness against non-linear harmonic distortion from the drivers, suitable for applying it to quiet and unoccupied rooms and the possibility of obtaining good results without a careful calibration of the system (Stan et al., 2002). It is created following Farina and Kite (Farina, 2000, Kite, 2004) recommendations:

$$s(t)_{Log_{chirp}} = \sin \left(\frac{2\pi f_1 T}{\ln \left(\frac{f_2}{f_1} \right)} \left(e^{\left[\frac{t}{T} \ln \left(\frac{f_2}{f_1} \right) \right]} - 1 \right) \right) \qquad (4.10)$$

Where:

$s(t)_{Log_{chirp}}$ = logarithmic sine sweep signal.

---

f_1 = initial frequency of the sweep. {Hz}

f_2 = end frequency of the sweep. {Hz}

T = duration of the sweep (Total time). {s}

A short gain ramp at the beginning and at the end of the signal is applied by multiplying the ramp and the generation of the sine sweep signal. The reason behind this is to avoid any start and stop transients[8]. This procedure avoids frequency artifacts caused when the sharp time transients are transformed to the frequency domain. This phenomenon may occur when a signal does not begin or end with a zero value. However, is recommended to design a sweep length as long as possible in order to obtain suitable resolution in low frequencies by having enough time to excite the room properly with the sound source.

4.2.4 Impulse response extraction method

The method of deconvolution is implemented in the frequency domain with the following expression according to Müller and Massarani (Müller and Massarani, 2001) and standardised in (Havelock et al., 2008):

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} \qquad \{adimensional\} \qquad (4.11)$$

Where:

$X(\omega) = \texttt{fft}(x(t))$ Spectrum of Stimulus equivalent to input of the system

$Y(\omega) = \texttt{fft}(y(t))$ Recorded Spectrum of signal equivalent to the output of the system

$H(\omega) = $ Transfer Function calculated in frequency domain

---

[8] http://www.jezwells.org/Computer_music_tools.html [Online accessed on Feb 8th 2012]

Therefore, the impulse response (*IR*) is found by applying the inverse fast Fourier transform (*IFFT*) to the transfer Function (*H*):

$$h(t) = \texttt{ifft}\big(H(\omega)\big) \qquad \{\text{adimensional}\} \qquad (4.12)$$

The chirp modelled in *Matlab* has a frequency response from 50 Hz to the current Nyquist limit (for this case is up to 96 kHz) with a period $T = 10$ s. These parameters are aimed to improve accuracy on the estimation of the impulse response at low frequencies where the source needs more time to build up the sound energy. It is normalised at -0.3 dBFS and then saved as a wav file. The reason is to limit the excursion on the studio monitor at low frequencies and therefore avoid any distortion artifacts in the measurements. It is then reproduced using Apple's *Logic Pro 9* (*DAW* software). The sound card balanced output is connected to the monitor speaker. Lastly, the recorded signals in the *DAW* were carefully calibrated to have at least headroom of 12 dB to avoid any chance of digital distortion on the recorded signals. The following **Figure 4.2** shows the practical acquisition of the data used to obtain the impulse responses outside of the semi-anechoic chamber, which is connected outside by the black box:
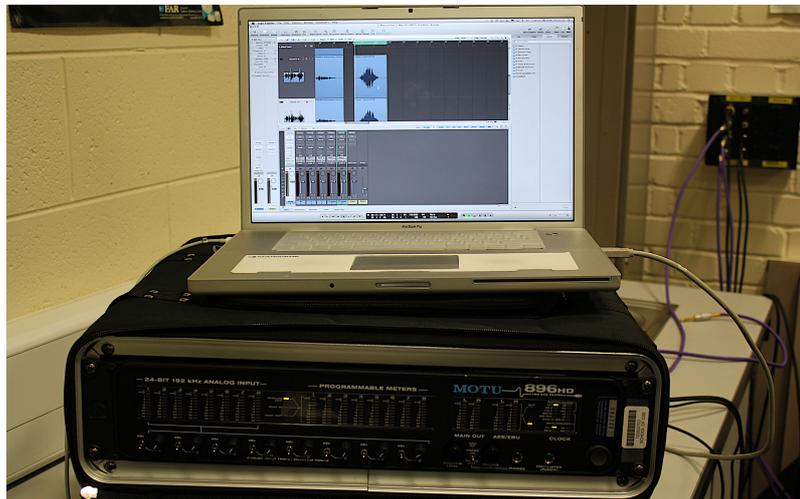
Figure 4.2: The first attempts to measure early reflections were made with a *Motu* 896HD sound card and the *DAW* software *Logic Pro* 9 running under a *MacBook Pro* laptop computer. In the monitor is possible to see two recordings of the room´s response to the stimulus signal.

4.3 Probe Configurations

The main goal of using sound-intensity techniques is to find an accurate and reliable method to measure early reflections, and it is crucial to get good estimations on a single point in space. A review of the existing products offered by the main companies that manufacture intensity probes found that G.R.A.S. Sound and Vibration company and Brüel and Kjaer company offer the most advanced intensity probes measurement system and analysing software in the market. Unfortunately, Brüel and Kjaer company does not offer a 3-D intensity probe in their catalogue anymore, just the one-dimensional version, the last 3-D version was found in (Fahy, 1995, de Bree et al., 1999). On the other hand, G.R.A.S. has two sound intensity probe models: the X-Y-Z Sound-intensity Probe Type 50VX[9] (2005b), which has an automated rotator mechanical system for a one-dimensional p-p probe. The other is the expensive state-of-the-art intensity probe, which is the Vector Intensity Probe Type 50VI[10] (2005a). In order to use this device a single alignment is needed along the axes. It does not need to rotate the probe anymore, because it has six phase-matched microphones. This accurate device was used by Tervo in his research (Tervo, 2012). For a high budget, this is the best option because it is practical and accurate. Nevertheless, the G.R.A.S. company offers it without any post-processing software.

4.3.1 The evolution of the one-dimensional p-p probe with cradle system

The core of an accurate measurement of intensity using a single p-p probe is to devise a method to accurately rotate the probe, maintaining its acoustic centre. The first attempts were made without a cradle system and the results were impractical and not satisfactory see **Figure 4.3**.

---

[9] http://www.gras.dk/50vx-intensity-probe.html [Accessed on May 4th 2013].
[10] http://www.gras.dk/media/docs/files/items/p/d/pd_50vi_ver_06_08_02.pdf [Accessed on Oct 21st 2013].
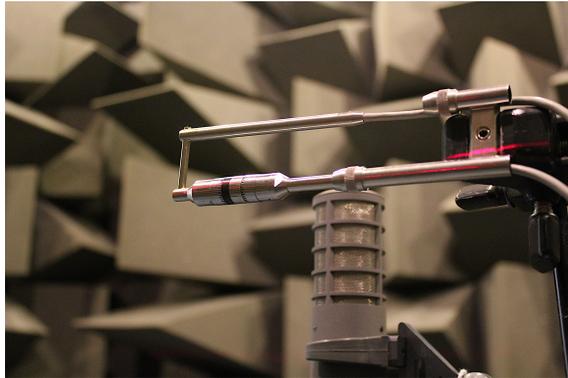
Figure 4.3: The first attempts to measure a 2-D sound intensity vector using the face-to-face configuration and the p-p probe with its existing holder.

The design of the custom-made rotation cradle had three development phases. The first one was accomplished in the mechanical workshop of the University of Salford see **Figure 4.4**. The original idea was to adapt the original holder provided by Brüel and Kjaer p-p intensity probe to a device that could rotate accurately 90º from its acoustic centre. This device was unreliable for the task. This fixture without a cradle was not accurate enough to obtain reliable results. The procedure used the laser cross to maintain the geometric centre of the probe, however the orthogonal rotation of the probe was time consuming and unreliable because there were too many degrees of freedom to control. In order to obtain better results, the use of *CAD/CAM* technology was needed.
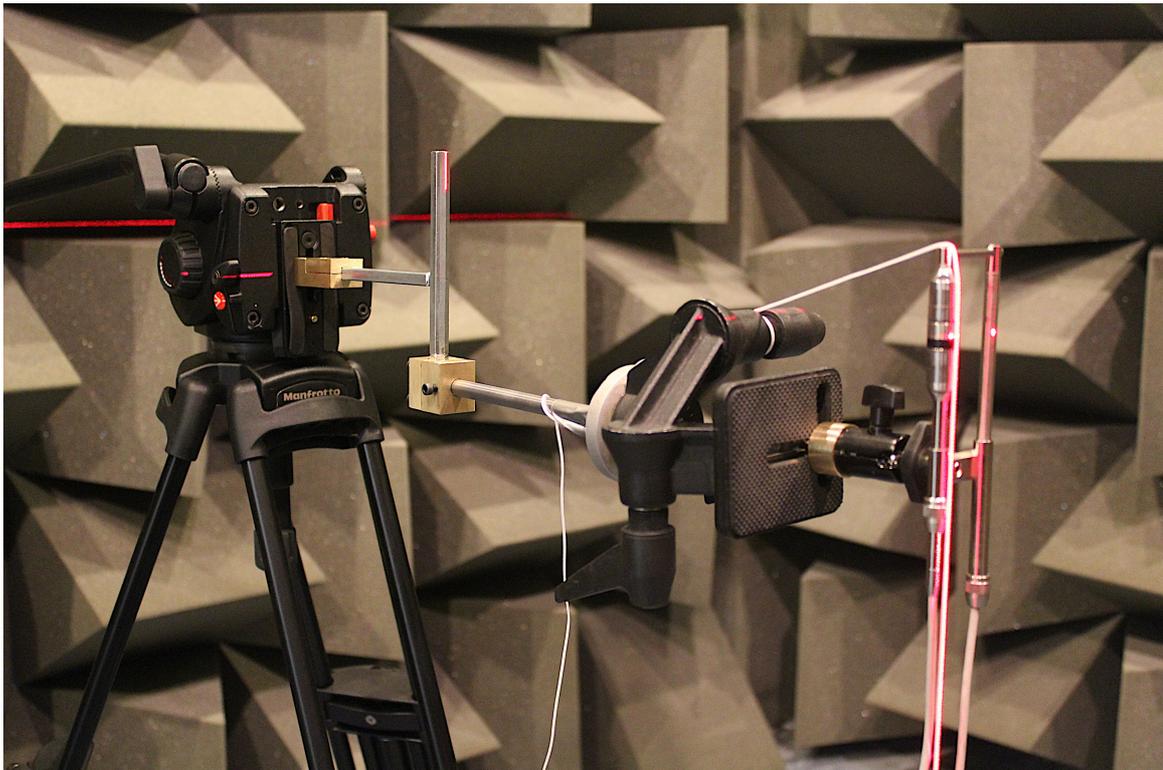
Figure 4.4: The first rotation cradle prototype was made in the mechanical workshop at the University of Salford, however it was not accurate enough to perform orthogonal rotations and it needed a *Manfrotto* head to align it vertically.

The design of the second prototype was done with the help of Minalum de México company. (Please refer to **Figure 4.5Figure 4.5:** Second prototype consisted on a "C" and a "L" shape machined in *CNC* Vertical Machine centre using high- grade hardened aluminium in Minalum de México company. The work piece was machined first by the upper side and later by the downside in order to make a single piece prototype without any bending or soldering to achieve high accuracy. The right photo of the same figure has a detail of the machining pieces "L" and "C" shape. The dark hole at the right is used to align the work piece after it is turned over by 180 degrees by accurately measuring its centre and align the *x*-axis. The next step was to use the wire cut *EDM* machine to cut two octagon shaped holes on the "L" shaped piece. The idea behind the octagons was to be able to rotate at 45º steps instead of 90º in order to resolve simultaneous arrivals. The diameter of the "C" and "L" pieces was chosen to be $\phi = 8.6 \text{ mm}$ in order to not disturb the acoustic field when the wavelength of the maximum audible frequency (

$\lambda/2 < 8.6$ mm $@ f = 20$ kHz) and it was planned to be round in order to create the same effect on all directions of incoming sound. The shape of this design was based on the custom cradle designed by (Abdou, 1994), (please refer to **Figure 4.6**). Unfortunately, this design did not comply with the required accuracy because the C and L shape were machined using two different procedures based on CNC. The octagonal components (male and female) were cut with a wire electro-discharge machine (EDM). Hence, a subestimation in the design provided no guarantee of obtaining an accurate rotation with the geometrical centre of the device. Additionally, the mounting system did not overlapped with the centre of the microphone base, which was a good option for locating the receiver when performing measurements at any given coordinate.

Figure 4.5: Second prototype consisted on a "C" and a "L" shape machined in CNC Vertical Machine centre using high- grade hardened aluminium in Minalum de México company.
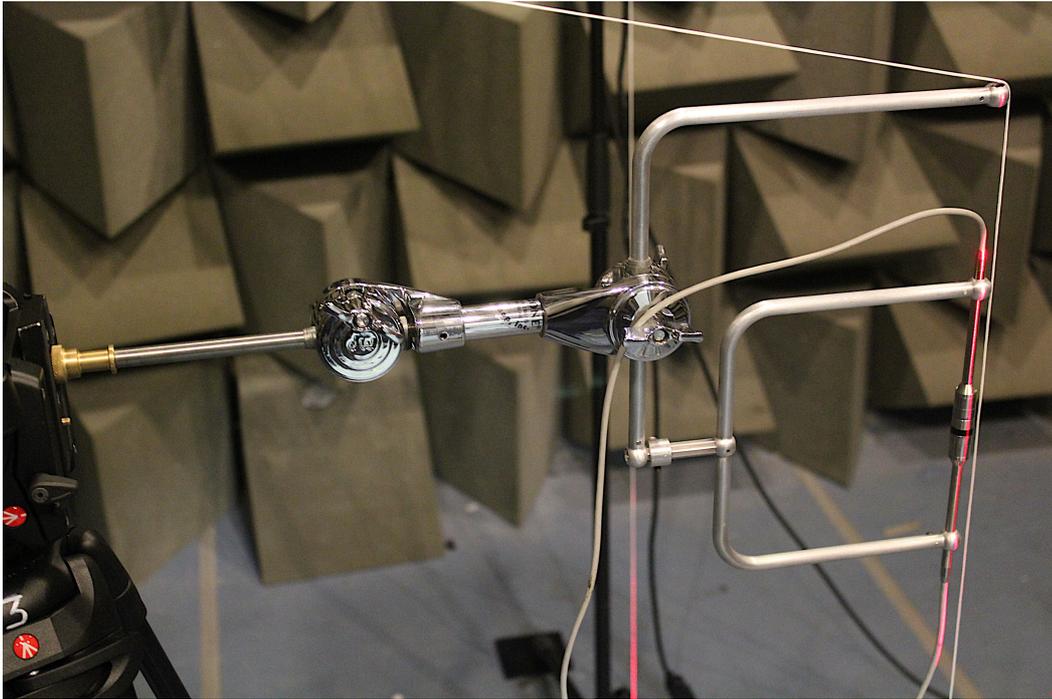
Figure 4.6: The second prototype for rotation cradle was tested for accuracy on the measurements of intensity. Unfortunately, it had a design problem where its geometric centre was not accurate for rotation. The best solution was to redesign a new prototype with an error proof rotation mechanism.

4.3.1.1 Design details of p-p intensity probe's custom-made rotation cradle

It was later concluded that the best way to design a rotation cradle was by only using Delcam *Power SHAPE* and *PowerMILL CAD-CAM* software and *CNC* and a vertical machining centre (*VMC*) in the facilities of Minalum de México company. The third attempt was done using only *CNC* machined parts, which were done with only one alignment on the piece to work. That realization guaranteed the precision required by using small machining steps and slow rate of removing of material. The third prototype introduced a number of new features that ensure accuracy on the setting rotation and repeatability on the measurements. The "O" carcass first was planned to be round in order to provide uniform scattering of sound and superior rigidity to be deformed with accidental misuse. In the last design, it ended as a squared shape to maximise the distance. The diameter of the shapes became slightly smaller than in the second prototype and the joints were round accordingly to the best acoustic fixtures.
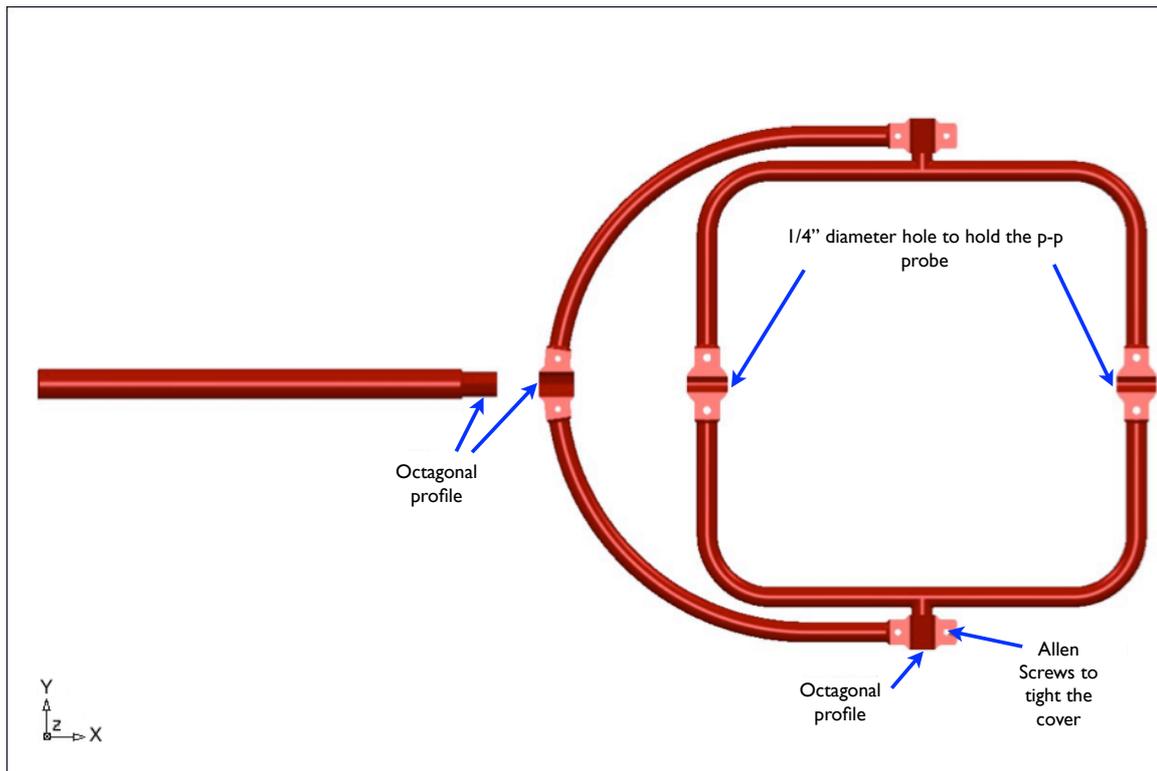
Figure 4.7: Detail of the assembly of the rotation cradle machined in *CNC* .
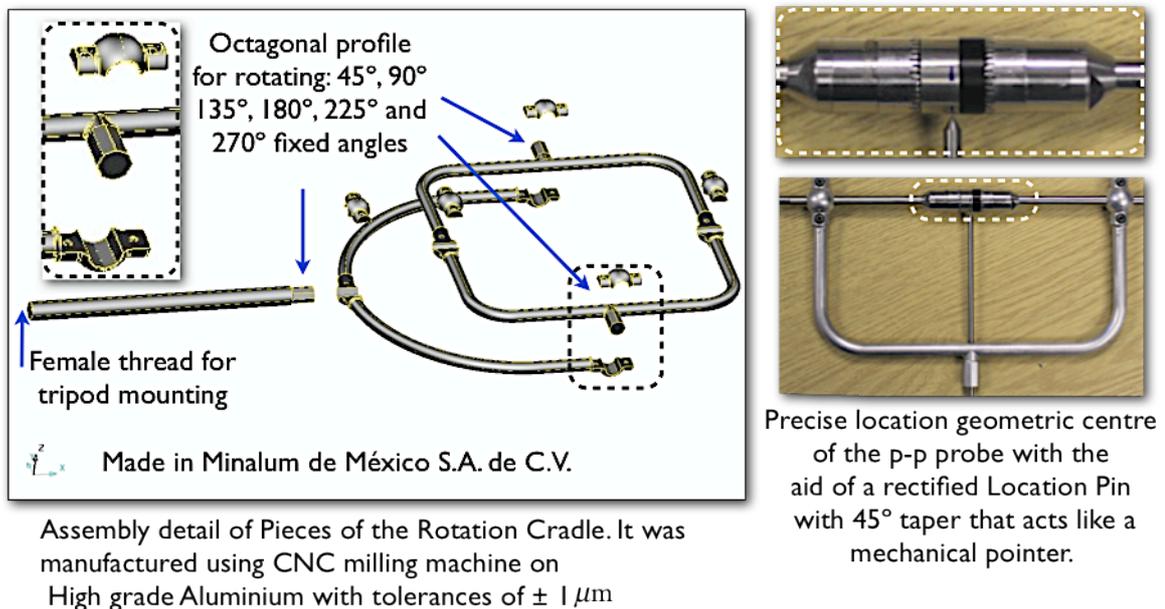
Figure 4.8: Assembly parts used in the rotation cradle using Allen 1/8" diameter type screws.

Figure 4.9: The third prototype of rotation cradle guaranteed the accuracy of the positioning of the face-to-face p-p probe configuration with the highest accuracy, the rotation of the p-p probe at 90 degrees.
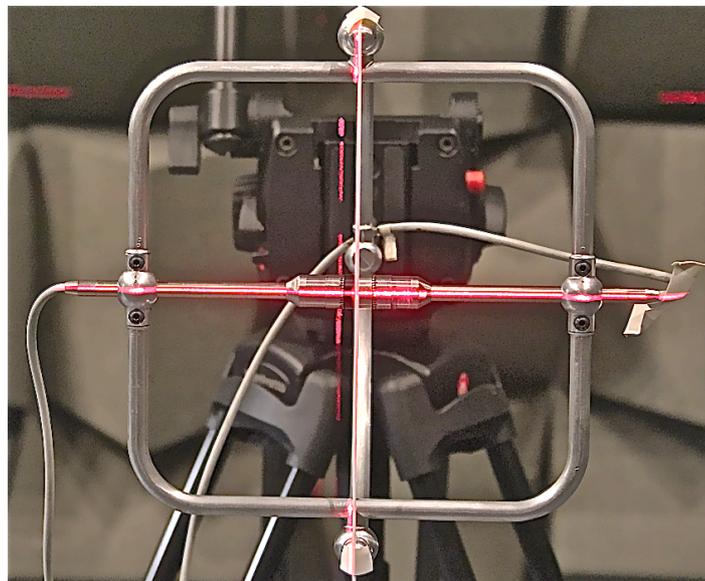
Figure 4.10: Custom rotation cradle aligned with its geometric centre with a laser cross and the plumb bob to align the vertical line with the laser to obtain an accuracy of about ± 1 mm.

4.3.1.2 The p-u *Microflown* custom-made aligner device

One of the problems encountered using a *Microflown* 3-D probe is how to measure the origin of the probe as a reference and how to assure that the device is accurately aligned to the measurement axis. The inner transducers of this device are quite small and in its carcass only has some printed points on its body to locate the 3-D axis. There is uncertainty of correct alignment and a custom-aligner is needed to ensure proper alignments. This device can be used in conjunction with the laser cross device to align the *Microflown* p-u probe. Since the company that manufactures the probe has not produced any alignment system, there was an opportunity to improve this measurement system.

The next design that was manufactured in Minalum de México company consists of a custom alignment system (please refer to **Figure 4.11**). This is a tool for the correct measurement of directional information when using the p-u *Microflown* USP intensity probe. One of the required features when performing a precision measurement aligning the probe with the orthogonal axis is to be able measure the accuracy of the setup at the exact centre position of the anemometer wires. This action is crucial because it determines the precise coordinate of the receiver. This task can be done with the aid of the dual laser cross system described in section 5.1.5 . This device helps to achieve the accurate position of the receiver with tolerances of ±1 mm. More importantly, it helps to align the axis of the probe along the axis used externally in the measurement setup. It works in the following manner:

The lower component of the system attaches the probe to the stand and aligns the centre of the probe with the centre of the microphone stand. The upper part of the system consists of a squared surface, which has a slit in the centre of it. The purpose of the slit is to make visible the anemometer wires and the pressure microphone at the centre of them. By securing this upper component with the Allen screw to the probe, both pieces can rotate together against the lower component. Doing this is possible to align the external axis along the external axis dictated by the laser cross system. Once this alignment procedure is finished, the squared upper component needs to be removed in order to

perform accurate acoustic directional measurements. It is crucial to avoid movement of the probe while performing this task to ensure that the coordinate measured with the laser cross is valid.
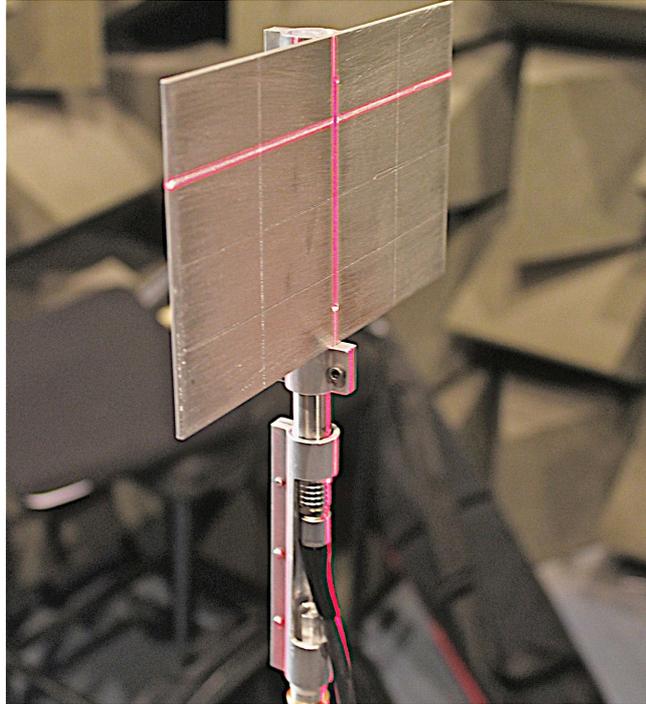
Figure 4.11: Custom co-lineal probe holder and alignment system for the p-u *Microflown* USP intensity probe.

4.3.2 The *Soundfield* microphone as a intensity probe

The *Soundfield* microphone is used to extract 3-D impulse responses (for details refer to (Craven et al., 2010)). It has four hyper-cardioid capsules, which generate the A-Format signal (refer to **Figure 4.12**). By applying a matrix algorithm within either a hardware unit or software implementation, it is possible to obtain the B-Format signal. It consists of four simultaneous signals: an omni-directional polar response signal (w), which represents the acoustic pressure (p) with unity gain. The other remaining three signals are orthogonal virtual figures of eight microphone polar responses centred at the origin. The *Soundfield* rack processor deliver the dipole Signals ($\vec{x}, \vec{y}, \vec{z}$) with a gain of $\sqrt{2}$ (+3 dB) referenced to the omni-directional signal w. This is standardised in order to give the same level as the signal w in free field conditions. The reason is to maximize the signal-

to-noise-ratio (SNR) for analogue recordings (Merimaa, 2006). In this analysis, the acoustic pressure (p) is taken from the value of the w channel i.e. ($p = w$). For further analysis in *Matlab*, the dipole signals ($\vec{u}_x$, $\vec{u}_y$, $\vec{u}_z$), which represent particle velocity vectors ($\vec{u}(n) = x(n)\hat{\mathbf{i}} + y(n)\hat{\mathbf{j}} + z(n)\hat{\mathbf{k}}$), are downscaled by dividing them by root of two (Farina and Ugolotti, 1998) as they were implemented in the post-processing.

This type of measurement requires a calibrated *Soundfield* microphone and a reliable sample and phase accurate soundcard connected to a silent portable computer. Measurement software performs the extraction of the impulse response of the four B-Format signals. (Refer to **Figure 4.12**). *Matlab* is used for further analysis a list of scripts is covered : *Matlab* scripts. The excitation source used is a *Genelec* 8030A studio monitor[11], which has a frequency response from 58 Hz to 20 kHz ± 2 dB.
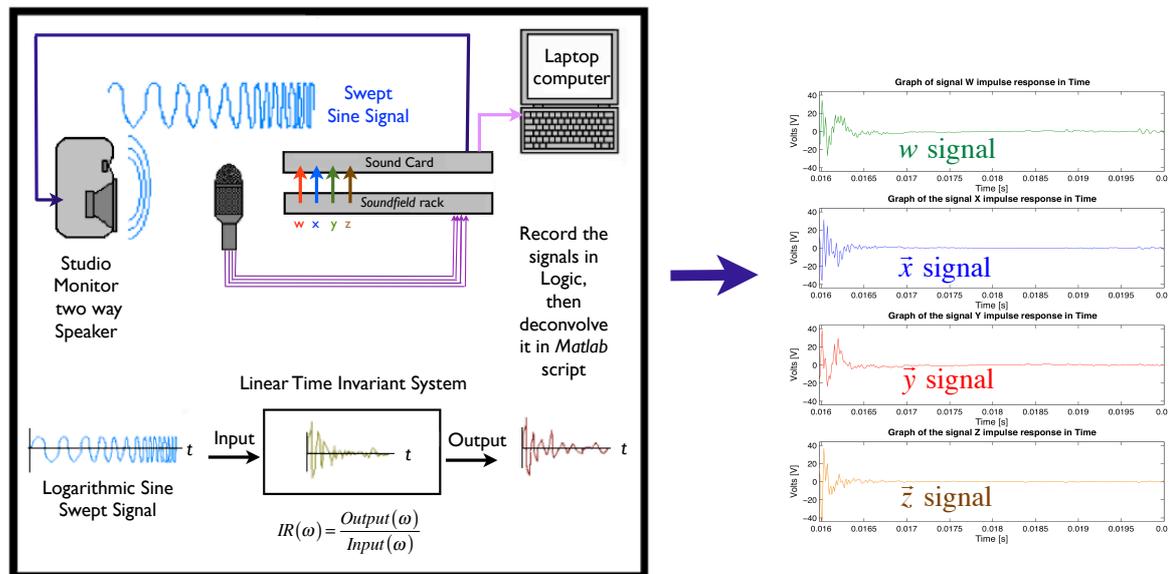
Figure 4.12: Impulse response (IR) measurement system using the *Soundfield* microphone model ST 350.

4.4 Processing and analysis of the data

For a practical implementation in a *Matlab* script, the instantaneous intensity measured with a p-p probe in time domain ($\vec{i}_{inst}$) needs to be implemented with a cumulative integration of the pressure, owing to the fact that the resulting integral must be a function of the time index (t) (Hanselman and Littlefield, 2005).

4.4.1 Time domain (TD) algorithm

Time domain is divided in smaller angular regions. By applying small time windows to the data, it is possible to analyse the early part of the decay, although not the late one. This is owing to the fact that reflections become too dense and are perceived as reverberation rather than isolated discrete reflections. Nevertheless, a statistical approach is still possible in the late decay zone using selective wider time windows to cover the entire reverberant tail. The following procedure analyses three-dimensional impulse responses (3-D IR) in time domain (TD):

1. Using non-overlapping windows, divide the array that represents the 3-D impulse response $\mathbf{h}_{3D}(t) = \begin{bmatrix} h_x(t) & h_y(t) & h_z(t) \end{bmatrix}$ in time chunks of equal duration i.e. ($\Delta t = 6$ ms.). Use non-overlapping windows on each orthogonal component.

2. Convert the time series to orthogonal components.

3. Convert the orthogonal components to dB scale normalizing with the maximum value of the 3-D impulse response (a time series array of size 3x1).

4. Calculate the magnitude of the signal and the angle of arrival in radians.

5. Plot the 2-D vectors of the active complex instantaneous intensity ($\vec{i}_{active\ complex}$) in polar coordinates.

6. Make a history of the evolution of the sound decay with an animated movie.

4.4.2 Data extraction algorithm

For the post-processing of the data, a combination of Time domain (TD) and short-time Fourier transform ($STFT$) was implemented. The m file named Extract_winMLS_signals_Corrected69LOG_BFormat_192kHz.m has the following parts:

1. Initializing program.

2. Constants.

3. Create a Sine Swept signal.

4. Reading raw data.

5. Band pass Filter design at 192 kHz.

6. Spectrogram for testing the quality of the signal before any processing.

7. Adding zero padding prior deconvolution.

8. Deconvolution done by spectrum division.

9. Integration for Intensity probe.

10. Scaling of the Intensity components in x and z-axis.

11. Conditioning to find the indexes of the max and min of the Intensity.

12. Logic to find the linear index of the maximum of the maximum and the minimum of the minimums.

13. Conditioning (Linear scale).

14. Logic for finding the abs maximum and its index without losing the original sample index.

15. Logic for finding the max_i and its index with 4 cases.

16. Logic to find the first arrival assuming x_ir and z_ir have same index.

17. Logic to find the beginning of the peak with the abs maximum peaks (Linear Scale).

18. Logic to create the chopped signal Time Domain.

19. Logic to create the chopped signal in log version.

20. Envelope Intensity (From Omoto 'Visualization of Sound Intensity' 2004).

21. Intensity by the Complex method of Heyser equations made with Heyser & Omoto 2002.

22. Intensity by the Complex method with the Envelope Omoto made with the Hilbert function (2002) .

23. Figures of the Complex intensity.

24. Figures.

25. Instantaneous Intensity of signal by finite differences method for the direct sound for the window 0-30 ms.

26. *STFT* method (adaptation to be more versatile with Fs).

27. Tic-toc measurement of time used to compute the program.

4.4.3 Short-time Fourier transform (Quiver plots and ERB)

The instantaneous intensity in frequency domain ($\vec{I}_{inst}$) results from the multiplication of the instantaneous value of the average pressure ($\bar{p}_{inst}$), calculated by using equation (3.18) on page 40 by the cumulative integral of the finite difference of pressures between the two microphones ($\vec{u}_{inst}$), and then apply the fast Fourier transform (fft) to it.

Merimaa applied the short-time Fourier transform ($STFT$) for analysing the directional impulse responses to characterise reflections (Merimaa et al., 2001, Peltonen et al., 2001, Merimaa and Pulkki, 2005, Merimaa, 2006). The method incorporated perceptual issues applied to the frequency content on the signals by using ERB. Later, Pulkki (Pulkki, 2007) and Ahonen (Ahonen et al., 2008) expanded its uses to teleconferencing under the *DirAC* method, although neither of them reported the errors in localisation of direct sound nor on errors in reflection estimation accuracy, as these errors are not important in teleconferencing. Enroth (Enroth, 2007) applied these techniques for auralisation of simulated acoustic impulse responses (IR) to be reproduced accurately in multiple channels using the *Soundfield* microphone, and combined the waveguide method to render synthetic signals.

Following the recommendation of Musha (Musha and Kumazawa, 2008) the short-time Fourier transform ($STFT$) method was used as one of the most accurate measurements for sound intensity ($\vec{i}_{inst}$) of transient signals. The $STFT$ analysis is used in the present work to obtain better estimation of reflection directionality than when it is computed in time domain (TD). It also helps to visualize the data in two dimensions at the same time. The downside of this approach is that once the window time resolution is set, the frequency resolution depends on the number of samples contained in the time window. Longer time windows enhance frequency resolution, but at the same time make a poorer time resolution (Smith, 1999, Lyons, 2002, Shin and Hammond, 2008, Tohyama and Koike, 2008). Using the highest sampling rate helps to minimise the frequency aliasing, this happens slightly at a low-frequency range, and becomes more noticeable at higher frequencies. This restriction means that a fine resolution in both domains cannot be achieved at the same time, which is a problem for the visualization of the same data because only one variable can be focused on at a given time. It is only with experience gained using these techniques that one can select the most suitable resolution.

On the other hand, the development of better resolution of the short-time Fourier transform method has been continued by Pihlajamäki (Pihlajamäki, 2009) by applying a multi-resolution based on longer time windows for low frequencies and shorter time windows for high frequencies. The more advanced transforms i.e. Wigner-Ville

transform, can help to overcome this limitation by offering better temporal and frequency resolution (Ville, 1948, Ferguson and Quinn, 1994, Tuncer and Friedlander, 2009). Currently, the most interesting approach found is called time-frequency reassignment (Flandrin et al., 2003). It is known that short-time Fourier transform ($STFT$) is still an approach that does not create artifacts because it is a linear transformation, whereas the Wigner-Ville transform still needs to be carefully used, since it is not a linear transformation.

This analysis involves two domains at the same instant. By using the short-time Fourier transform ($STFT$), it can map the behaviour of a variable in time and frequency. Prior to applying the windowing process, each complete component of the B-Format file (i.e. *w*, *x*, *y*, *z*) is conditioned by adding zeros with a length of *N/2* samples before, and adding zeros with a length of *N* samples after the signal. The reason for this action is because when the input is windowed, the first window can start at sample number one with its window's maximum value located at sample *N/2*. The next step is to divide each of the B-Format impulse responses by a chosen number of time windows of length *m* and applying the fast Fourier transform to the product of each of the B-Format signals with a Hann window (w_{Hann}). Each fast Fourier transform (`fft`) has a length of *N* points. Subsequently, the next window will start without any gain. Hence, any information in the first *N/2* samples prior to the real beginning of the signal will be attenuated according to the gain dictated by the shape of the first window. Applying zero padding to the first sample of the real signal ensures that the signal will be not be attenuated by the first window. In a similar manner, the next samples will always be located within two windows, which have a maximum gain and a minimum gain as a crossover point.

The time windows need to overlap each other in order to create a smooth transition between each window. Each time a window is zero-padded before and after it with half of the length of the time window (*N*). This is done to obtain the first sample multiplied with a window, which is at the maximum value to avoid spectral aliasing. The practical overlapping percentage applied to each Hann window is 50 %. (see figure **Figure 1.1**)
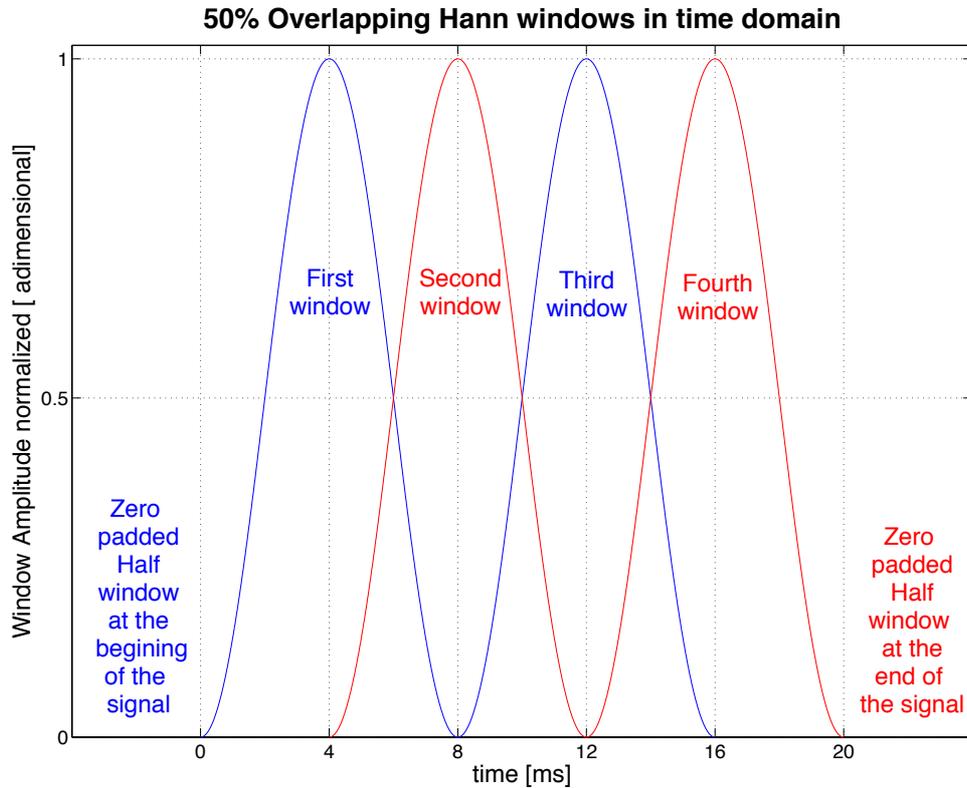
Figure 4.13: Example of a 50% Overlapping of Hann windows in time domain (TD). Note that the first window is zero-padded half a length of a window before and after the end of the train of overlapped windows with half of the length of the time window.

In order to create each slice of time, successive `fft` are applied to each component:

$$W(k) = FFT\{w(n) \cdot w_{Hann}(n)\} \tag{4.13}$$

$$\vec{X}(k) = FFT\{\vec{x}(n) \cdot w_{Hann}(n)\} \tag{4.14}$$

$$\vec{Y}(k) = FFT\{\vec{y}(n) \cdot w_{Hann}(n)\} \tag{4.15}$$

$$\vec{Z}(k) = FFT\{\vec{z}(n) \cdot w_{Hann}(n)\} \tag{4.16}$$

Where k is the frequency index calculated by the `fft`. The three-dimensional particle velocity vector in the *STFT* domain is created by applying the short-time Fourier transform (*STFT*) to each orthogonal component by adding the three components:

$$\vec{U}(n,k) = [X(n,k)\hat{\mathbf{i}} + Y(n,k)\hat{\mathbf{j}} + Z(n,k)\hat{\mathbf{k}}] \tag{4.17}$$

Where n is the time index for each slice of time.
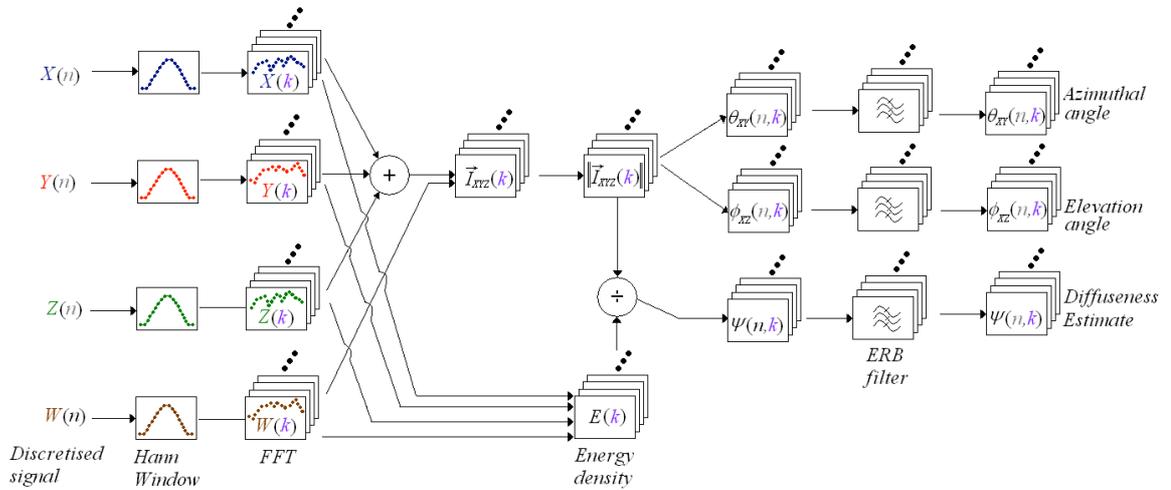
Figure 4.14: Diagram of the short-time Fourier transform (*STFT*) method applied to calculate instantaneous intensity ($\vec{I}_{inst}$) and diffuseness estimate (ψ) after Merimaa.

4.4.4 Peak detection in time domain (polar plots) with Envelope of signal $\overline{s}(t)$

After a validation of the measurement system is performed in simpler cases, it is possible to adapt the method to measure three-dimensional impulse responses (3-D *IR*) of real rooms. Because of the complex patterns generated under these cases, a peak detector algorithm needs to be used with the complex instantaneous intensity ($\vec{i}_{complex}$), or alternatively, apply the envelope of intensity ($\overline{i}$) on each x, y and z-axis components.

The best way to detect reflections is to use the envelope of the magnitude of the signal $\overline{s}(t)$ on each axis component, because the rapid oscillations of the decay of sound can be

removed in that way and the envelope can show the relevant information (Thrane et al., 1997). It is a good option for preserving the important reflections without removing important details on the impulse response (Kuttruff, 2000). Alternatively, the detection of reflections using microphone arrays tends to add noise to the signal depending on the distance between the microphones. Tervo has studied these cases in (Tervo, 2009) using an orthogonal array of 6 inner microphones and 6 outer microphones called the TKK-3-D microphone array, which is an open 3-D spherical microphone array (Tervo et al., 2011).

The best algorithms for peak detection take into account the effect of noisy signals where the peaks and valleys need to be detected (Billauer, 2011). The strategy defines a peak as the highest point between lower areas called "valleys." A condition to find a peak is to realise that there are lower points before and after the peak. Since a sound wave usually has positive and negative excursions as a natural form of propagation of a wave in the air, it is useful to combine the peak detector with the envelope of the signal in order to detect only relevant information. Then, select the relevant reflections and neglect the false peaks that are below a threshold determined by the perception of reflections, which is about a value of -30 dBFS relative to the direct sound, which is set to 0 dBFS (Begault et al., 2001, Begault et al., 2004, Schlemmer, 2006, Reed and Maher, 2009). The maximum value was set according to direct sound value, and it is very useful to compare it with reflections.

4.5 Summary

This section describes the methodology followed on the acoustic measurements with descriptions of the experiments undertaken. Implementing the measuring system explains the use of three-dimensional impulse responses 3-D *IR* using the *Soundfield* microphone model 350, the p-p probe, digitalizing the data and post-processing it using sound intensity using time and frequency methods. The first approach was to determine the accuracy of the measurement system by simplifying the scenario to only one reflection. Subsequently, the multiple reflection scenarios are described and introduce the simultaneous reflection problem and how it was tackled. The discussion continues with

the techniques used to minimise the uncertainty of the results. The use of damping vibration devices and laser cross is discussed in order to ensure the accuracy of the positions measured. The approach for determination of uncertainties of the data was obtained with a sample of 10 measurements to compute the uncertainty of the error. The mean error ($\bar{\varepsilon}$) and weighted average of standard deviation mean error ($\bar{\sigma}_{\bar{\varepsilon}}$) were used to quantify the accuracy of the measurement system. The percentage error ($\varepsilon_{\%}$) is also covered in order to make a decision of which parameters to report.

The novel contributions of the author are in the design of the custom-made rotation cradle, and the post-processing and analysis of the measured data in order to use intensity to characterise the strength, spatial and temporal attributes of the early reflections in small rooms.

Chapter 5 : Acoustic probe configurations

The details of the acoustic probe configurations used in the measurement environment are described in this section. They are comprised of two main tasks: a) laboratory tests to validate the method and b) practical applications of this method for measuring a small room, which exhibits multiple reflections. The details of the experiments are described in this section. A description of the measurement conditions is followed. It is comprised of the specifications of the semi-anechoic chamber, the measurement position calculation, the influence of the directivity of the sound source, the use of damping devices to minimize contamination by vibration in the measurements, the multiple laser cross system to ensure accuracy on positioning of receiver and source experiment designs useful for measuring early reflections. The last section of the chapter covers the application of the theory used to measure the incoming sound in a single reflection scenario and the details of uncertainty treatment used in order to confirm which is the practical extent of this method.

5.1 Description of measurement environment

5.1.1 Semi-anechoic chamber acoustic specifications

In order to have less noise in the impulse response, the one-reflection case was implemented in the semi-anechoic room G-35 in the Acoustic Laboratory of the University of Salford. The Background noise level is reported as low as 3.8 dBA, the inner working dimensions are 4.2 m $\times$ 3.3 m $\times$ 3.0 m and its cut-off frequency ($f_{\text{cut-off}}$) is 250 Hz. Therefore, for this experiment, the practical frequency working range is from 250 Hz up to around 10 kHz, which is the limiting frequency using the p-p probe using a 12 mm spacer and the *Soundfield* microphone.

5.1.2 Measurement positions and angles (real values)

For detailed information of the calculation's procedure for the geometric model is referred in the following steps:

a) For the determination of the real angles of arrival of direct sound (θ_{s_i}) and for the distance over *x*-axis (x_i which correspond to the ith position P_i), please refer to equation 4.2 in page 63. The results of this calculation are the input data and it is depicted in **Table 5.1** and **Table 5.2**.

b) For the calculation of the angles of arrival of first reflection (θ_{m_i}), please refer to 4.4. in page 64, **Table 5.3** and **Figure 5.9**.

5.1.3 Monitor speaker influence

Tervo et al. (Tervo et al., 2009, Tervo, 2012) recommends the use of a highly directional loudspeaker for the detection of the early reflections, since this method approaches the ray tracing method and analyses only a discrete number of reflections depending on the orientation of the source. The downside of this approach is that it is extremely time-consuming, because it requires performing several measurements rotating the highly directional speaker over the horizontal plane to cover steps of at least of 10º. The same case occurs in the vertical plane, however the highly directional speaker is more difficult to rotate in that direction.

In contrast, the same measurement can be done with one omni-directional or partially directional (source studio monitor) in only one measurement if a broadband monitor is used. That approach was partially fulfilled by this study by using a near field monitor *Genelec* 8030A[12]. It has a reasonably wide directivity in the horizontal plane (*x-y*) and a

---

[12] www.genelec.com/documents/datasheets/DS8030a.pdf [Online accessed on April 25th 2013]

very tight directivity in the vertical plane (x-z)[13]. It has a waveguide in the tweeter to limit the spreading in the vertical plane to some extent. Therefore, it may limit the *SPL* level in the vertical plane, as it tends to show a more uniform *SPL* in the horizontal plane regarding the angle sustained in the *x-z* plane. The configuration used for the measurements was planed to cover a limit angle range constrained to the internal workable dimensions of the semi-anechoic chamber.

However, the two-driver monitor is contributing to the difference path from the woofer and the tweeter, which is frequency dependent. The active crossover divides the signal, and because the drivers are not concentric, there is a certain delay on the time of flight (*TOF*), which additionally corresponds to a difference of direction of arrival (*DOA*) for certain frequencies covered by a different driver. Moreover, the transient response of both drivers is different because of the mass of each driver and the efficiency of the magnet and the power required to operate. Therefore, the woofer tends to add another delay to the frequencies that it covers. The result is that the measured impulse response of a room will present a phase difference, which is a function of the crossover frequency ($f_{crossover}$). One way to overcome this problem is to use the acoustic centre of the monitor for the estimation of the sound source real angle of arrival, but this will not be accurate for all frequencies (see **Figure 5.1**). On the other hand, to improve localization of reflections, Rechenberger (Rechenberger, 2009) recommends using a minimum phase transformation applied to the impulse response (*IR*). The non-flat frequency response of the speaker similarly can be filtered with an inverse counterpart and then multiplied to the spectrum of the impulse response (`fft` (*IR*)). This process may be applied to acquire an ideal flat response from the speaker. Conversely, these processes introduce delays to the signal, which may affect the time domain analysis to some extent. For that reason, it was eluded in the post-processing of the analysis of reflections in this thesis.

---

[13] http://www.princeton.edu/3D3A/Publications//Speaker_Directivity_Data.pdf [Online accessed on April 25th 2013]

Figure 5.1: Acoustic centre of the near field monitor *Genelec* 8030A used in the measurements located by the laser pointer located at tangent to the woofer at a height of 19 cm from the Iso-Pod™ table stand of the monitor and valid from a distance larger than 0.7 m (2007).

A typical speaker has a frequency dependant directivity, which varies between omni-directional and directional. At low frequencies, it tends to show an omni-directional directivity, nevertheless the directivity increases with mid and becomes narrower at high frequencies. Ideally, the best source directivity for measuring all the reflections in a room should be omni-directional. One way to obtain that is to attach a cone to a speaker as shown in (Li, 1995, Miura et al., 2010). Using it in this way can correct the deviations at low and high frequencies. The downside is that it needs to have a single cone for each driver to cover different frequency ranges, which means that the measurement of an impulse response (*IR*) needs to be carefully done in several stages if a single point source is needed. An alternative way to manufacture a 'point source' is by using a wide band compressor driver coupled with a tube termination of a certain length. Its frequency response is not flat because of the tube resonances impinged to the frequency response. However, the advantage is to be able to easily determine the acoustic centre of the source as it is located with the two laser cross system as can be seen in **Figure 5.2**. Unfortunately, the power radiated by this device is compromised with the distortion of

the driver and its inefficient acoustic radiation, because of the resonances of the coupled tube and the generated wave reflections by the sharp ending of the tube. This idea was partially tested with a custom simple prototype, composed of a compressor driver and a pipe, which was not used for results in this thesis, although it is considered for future work (please refer to **Figure 5.2** on the left side). For professional results, the measurements should be done with the Brüel and Kjaer *OmniSource* Sound Source Type 4295[14] to avoid the coloration of the sound by the resonances of the tube and to manage enough acoustic power to obtain good a signal-to-noise-ratio (*SNR*).
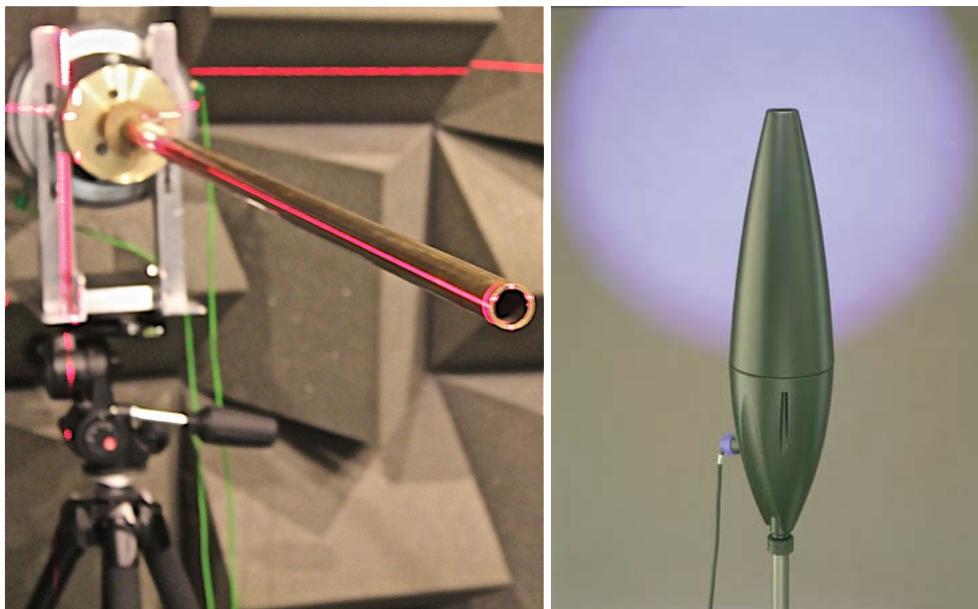
Figure 5.2: Examples of two omni-directional sound sources. The left one was a custom-made at the University of Salford, and the one on the right is the Brüel and Kjaer *OmniSource* Sound Source Type 4295, which compensates the resonances of the tube by varying the diameter of the tube.

---

[14] http://www.bksv.com/doc/bp1689.pdf [Online accessed on May 25th 2013]

5.1.4 Damping vibration on source and receiver

The effects of random low-frequency noise introduced by surface borne vibrations were taken into account. A vibration-dampening device decouples the source stand. It consists of a rubber mat especially designed for decoupling speakers.

The stand used to mount the microphone array was decoupled using the same approach. Care was taken to follow the recommendations of using microphone stands and microphones with angle of zero degree between them in order to minimise the sound field disturbance (Svend, 1985). All experiments were conducted without intervention of a person inside the measurement room in order to maintain the maximum signal-to-noise-ratio found in the semi-anechoic chamber. The difference between the measurements with the vibration-dampening device and the tripod without the device were not measured in order to obtain an objective comparison. However, the accuracy of the results improved significantly, as it can be visually compared with the early measurements. In the early measurement, it can be noticed that there are larger estimation errors in the angle of arrival of direct sound and first reflection especially at high frequencies and low frequencies (please compare **Figure 5.3** and **Figure 6.56**).
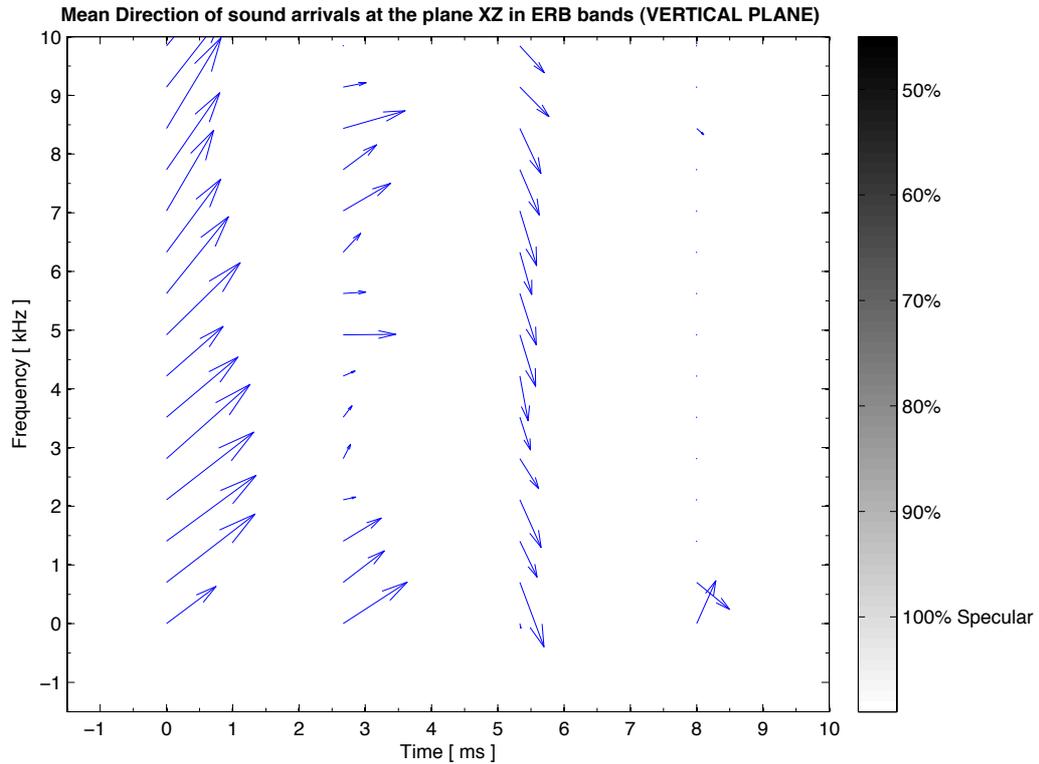
Figure 5.3: Example of measurement of p-p probe using the *STFT* made without the damping vibration on the source and the receiver. The measurement position P_1 is also called the calibration position at 45º for the direct sound.

5.1.5 Laser cross system to define the real spatial coordinates

To ensure that all the variables were tightly controlled, a number of measures were applied. It was necessary to use laser pointers with a self-levelling system, which can deliver reliable horizontal and vertical lines (*Black and Decker*, *Lasercross* LZR6TP[15]). The laser used is a consumer level device that has an internal pendulum system where the dual orthogonal laser head is self-adjusted by gravity whenever it is moved. The reported accuracy is +/- 1.5 mm/m. It was found that in order to make reliable measurements of the receiver position, it was necessary to have two orthogonal laser pointers to keep the alignment of the microphone and the geometric centre of the speaker under a co-lineal

---

[15] http://www.blackanddecker.co.uk/powertools/productdetails/catno/LZR6TP/info/uses/
[Online accessed on May 11th 2013]

arrangement. The estimation of the 'real angle' of the reflection (θ_{m_i}) was done with the measurements on the x-axis. For the floor markers, a couple of two single line laser pointers were used to ensure the accuracy desired. One was used to read the coordinate of the x-axis while the other was used to maintain the alignment of the microphone tripod, which was mounted on a wooden table with four attached wheel carts (refer to **Figure 5.5**)

Additionally, a rope was attached to the source and the receiver. At the end of both, a plumb bob was set almost touching the measurement ruler. It was fixed to the floor with tape (refer to **Figure 5.6**). The sharp end of the plumb bob was used as the measurement point. The intersection of the vertical point with the horizontal was set as the origin for the linear measurements in x-axis. The microphone and the origin of the vertical line from the speaker were used to create all the coordinates used for the positions measured. With these tools, the accuracy of the measurements was tightened up to an uncertainty of ±2 mm, since the width of the laser beam is around 1 mm (see **Figure 5.6**). It was sufficient for the purpose of the experiment, because the minimum change in each step angle consisting in one-degree position was set to a distance over the ruler of 49 mm. Consequently, the uncertainty of the measurement represents only ±0.04° of a degree (±4% variability), which is quite acceptable for the measurement in this scenario. The main limitation of this measurement system was human error involved in the measurements, which can become a cumulative error at the end of the sequence of measurements given that the rulers were not long enough to cover the entire set of distances. In order to control this problem, a concatenation of rulers was used. The parallax measurement error was compensated with the use of a small railway cart (refer to **Figure 5.5**).
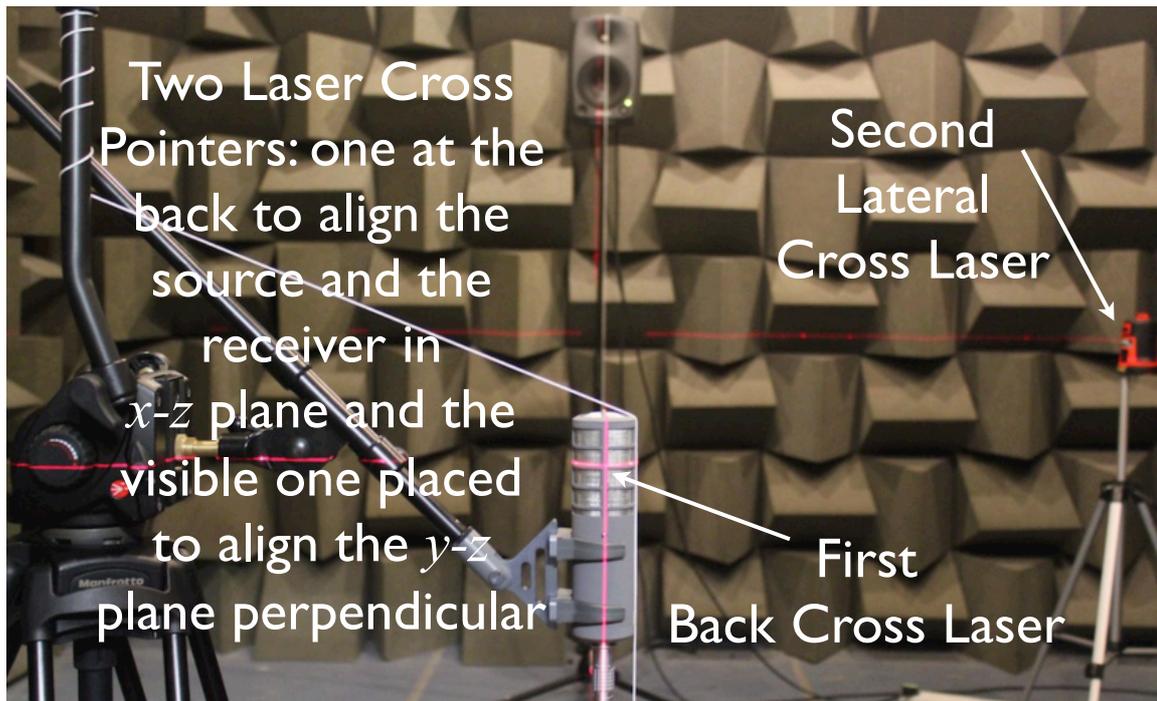
Figure 5.4: Measurement of single reflections captured by the *Soundfield* microphone model ST 350.

Figure 5.5: The laser pointer controlled in a railway trolley cart and the perpendicular steel ruler serving as a guide with the plum bob pointing at the ruler.
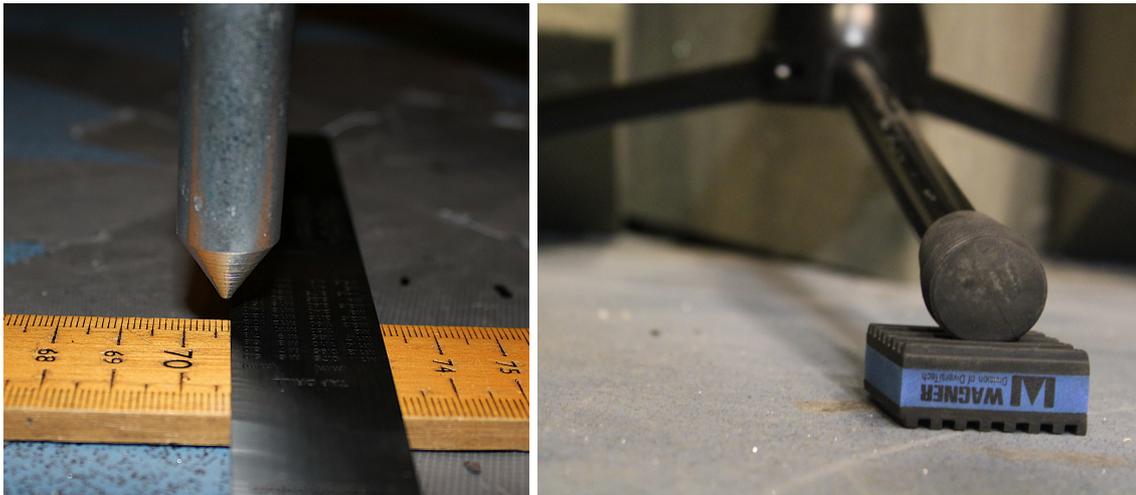
Figure 5.6: Measurement technique used with the plum bob (left) and the vibration-dampening device that decouples the microphone tripod (right).

The coordinate for position P_1 is calculated as $x_1 = \cos(45°) = 0.707 \text{ m}$, which is approximated as ≈ 70.7 cm. The measurement is performed with the plumb bob and the laser pointer technique. The origin is also verified at the beginning of the ruler. Please refer to **Figure 5.6**.

Throughout the research process, post-processing of the signals and new techniques to measure with more control has been also carried out. This has been focused on improving the accuracy of the linear measurements with angles by using multiple laser pointers to maintain the level of precision of measurements down to ±2 mm.

5.2 Laboratory tests for a single reflection in the semi-anechoic chamber

To assess the accuracy of a number of methods for measuring individual and aggregated room reflections, a series of scenarios were investigated. This subsection will now discuss the methodology applied in each scenario.

5.2.1 Single reflection scenario

A logical research method pursues from the simple to the complex situations. In order to deal with the simplest case, a single reflection case was chosen. This was the natural choice, since it provides a simpler response without interaction of many reflections to extract from the time series. A single reflection case for the impulse response (IR) was implemented in the semi-anechoic room at the acoustic laboratory of the University of Salford. The background noise level is reported as 3.8 dBA, and the inner working dimensions are 4.2 m $\times$ 3.3 m $\times$ 3.0 m. The usable frequency for this room is above 250 Hz.

The signals were acquired at 24 bits 192 kHz with three different step angles. From position P_1 to P_6, the angular step was one degree, from position P_6 to P_{11} the angular step was two degrees, and from position P_{11} to P_{14} it was five degrees (See **Figure 5.8**).

The first position is for calibration of the system. By choosing P_1 to deliver an angle of 45° for the direct sound (θ_s), the measurement system is tested on its capability to deliver symmetrical values on orthogonal axes. The last position P_{14} also gives a calibration value near 45° for the first reflection arrival angle ($\theta_{m_{14}} = 45.067^\circ$) (refer to **Table 5.2** and **Table 5.3**). This arrangement is useful to analyse the minimum resolution of the measurement system by giving the possibility to create three different analyses in fixed steps with five positions for each of the three sets. The difference of this method against the methods used by other workers is that this method focuses in a very important effect, which has been overlooked. This is the effect of varying the distance of the source and receiver and how it impacts in the accuracy of the angle of arrival of the direct and first reflection. Most of the studies had been focused in a horizontal rotation of the source and that maintain the same radial distance while the angle is changed. Therefore, the only variable is the rotation angle, and not the angle of the first reflection is kept constant. That assumption may be good to characterise the capabilities of an intensity probe in order to generate a graph with the capabilities for rotation of the direct sound across a

specific plane. Nevertheless, these approaches are not reporting anything about the effect of varying the distance while there are two variables on direct sound and first reflection. The study presented in this thesis is more complete because it shows the real behaviour of a intensity probe against different patterns of direct sound and first reflections as it may occur in a real room. The method considered the repeatability of the individual 14 positions ($P_1 - P_{14}$) and the sample size chosen was $N = 10$ samples per position. The Student's t-distribution was used to correct the standard deviation calculations by multiplying the standard deviations by a factor of 1.833 corresponding to match the N-1 = 9 degrees of freedom and the probability of 0.95. This is found in the corresponding table (Navidi, 2006). For each position, an estimation of the angle of arrival of direct sound (θ_{s_i}) and first reflection (θ_{m_i}) was performed by two main methods herein named: a) time domain method (TD) and b) short-time Fourier transform method ($STFT$).

The stimulus chosen was a logarithmic sine sweep signal generated using Farina's method (Farina, 2000) with a frequency range of 50 Hz up to 96 kHz. The reason for this is because the Nyquist limit is imposed by the chosen sample frequency of 192 kHz. Initially, the original recorded signal was not band-pass filtered, the reason was to use the untreated signal, keeping its bandwidth and avoid frequency artifacts created by filtering the signal.[16]

The equipment for the measurement system consists of the following items:

a) *Soundfield* microphone model ST 350 and its portable signal conditioner,

b) p-p intensity probe with G.R.A.S. phase-matched capsules and a Brüel and Kjaer ¼" internal preamplifier and a Norsonic signal conditioner model Frontend type 335,

c) *RME Fireface UFX* soundcard sampling at 192 kHz, with a bit depth of 24 bit using *DAW Pro Tools* 9 which can record at internally at 32 bits,

---

[16] http://www.rme-audio.de/en_products_fireface_ufx.php [Online accessed on April 25th 2013]

d) Several scripts of *Matlab* written by the author were used to acquire and condition the data (For more information please refer to Appendix A: *Matlab* scripts and physically there is a photo of the equipment in **Figure 5.7**).

Figure 5.7: The equipment used for the last measurements. The photo of the left side shows the preamp for the *Soundfield* microphone ST350 and the Norsonic signal conditioner type 336 for the one-dimensional p-p probe. The photo of the right shows the *RME Fireface UFX* sound card which uses digital preamps connected to a *DAW Pro Tools* 9 running over a *MacBook Pro*.

The measurement method consists of capturing both the sound pressure (p) and particle velocity ($\vec{u}$) in the 3 orthogonal directions at a single position in space. The acquired data was examined using the *STFT* technique, with 10 ms Hann windows and 50% overlap. For each *STFT,* the pressure and particle velocity were used to extract the direction of the sound energy flow $\left|\vec{i}\right|e^{\theta}$. The technique allows measurement of reflections with the highest accuracy compared with prior attempts in both horizontal and vertical planes.

•14 measurement positions were tested in steps of 1°, 2° and 5° covering an angle range of 30°.

Source Position

Receiver Positions

14   13   12   11 10   9   8   7   6 5 4 3 2 1

z

x

Rigid Floor in the semi-anechoic chamber
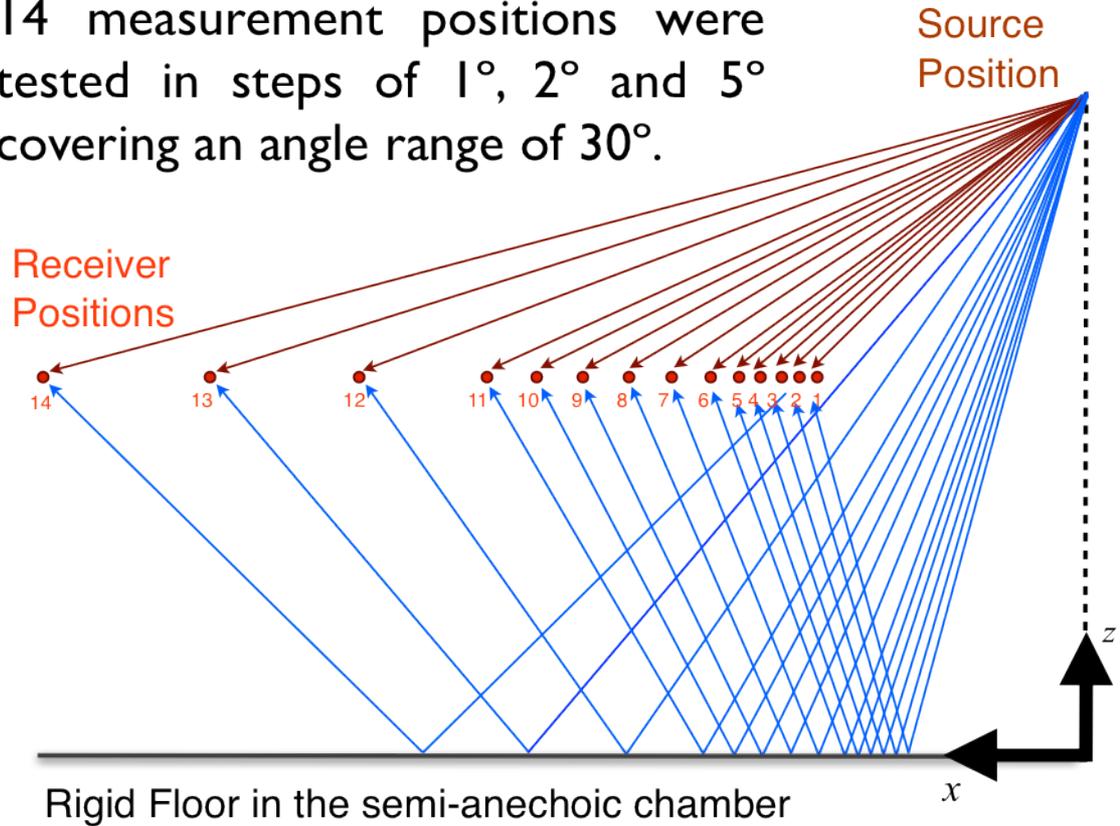
Figure 5.8: Sketch of the 14 different positions tested, where the source is static and the receiver is moved to measure different angles of arrival of direct sound (θ_{s_i}) and first reflection (θ_{m_i}). An image source model is used to estimate the 'real angle' of the first reflection (φ_i).

P_1	0.707 m	P_8	1.177 m
P_2	0.758 m	P_9	1.289 m
P_3	0.807 m	P_{10}	1.406 m
P_4	0.858 m	P_{11}	1.526 m
P_5	0.909 m	P_{12}	1.852 m
P_6	0.961 m	P_{13}	2.219 m
P_7	1.068 m	P_{14}	2.645 m

Table 5.1: Calculated distance over x-axis in meters with positions P_1 to P_{14}.

θ_{s_1}	45.000º	θ_{s_8}	31.002º
θ_{s_2}	43.052º	θ_{s_9}	28.742º
θ_{s_3}	41.218º	$\theta_{s_{10}}$	26.702º
θ_{s_4}	39.491º	$\theta_{s_{11}}$	24.853º
θ_{s_5}	37.865º	$\theta_{s_{12}}$	20.898º
θ_{s_6}	36.331º	$\theta_{s_{13}}$	17.671º
θ_{s_7}	33.517º	$\theta_{s_{14}}$	14.966º

Table 5.2: Angles of arrival of direct sound (θ_{s_i}) with positions P_1 to P_{14} from the source.

θ_{m_1}	-75.067º	θ_{m_8}	-66.067º
θ_{m_2}	-74.067º	θ_{m_9}	-64.067º
θ_{m_3}	-73.067º	$\theta_{m_{10}}$	-62.067º
θ_{m_4}	-72.067º	$\theta_{m_{11}}$	-60.067º
θ_{m_5}	-71.067º	$\theta_{m_{12}}$	-55.067º
θ_{m_6}	-70.067º	$\theta_{m_{13}}$	-50.067º
θ_{m_7}	-68.067º	$\theta_{m_{14}}$	-45.067º

Table 5.3: Angles of arrival of first reflection (θ_{m_i}) with positions P_1 to P_{14} from the source.
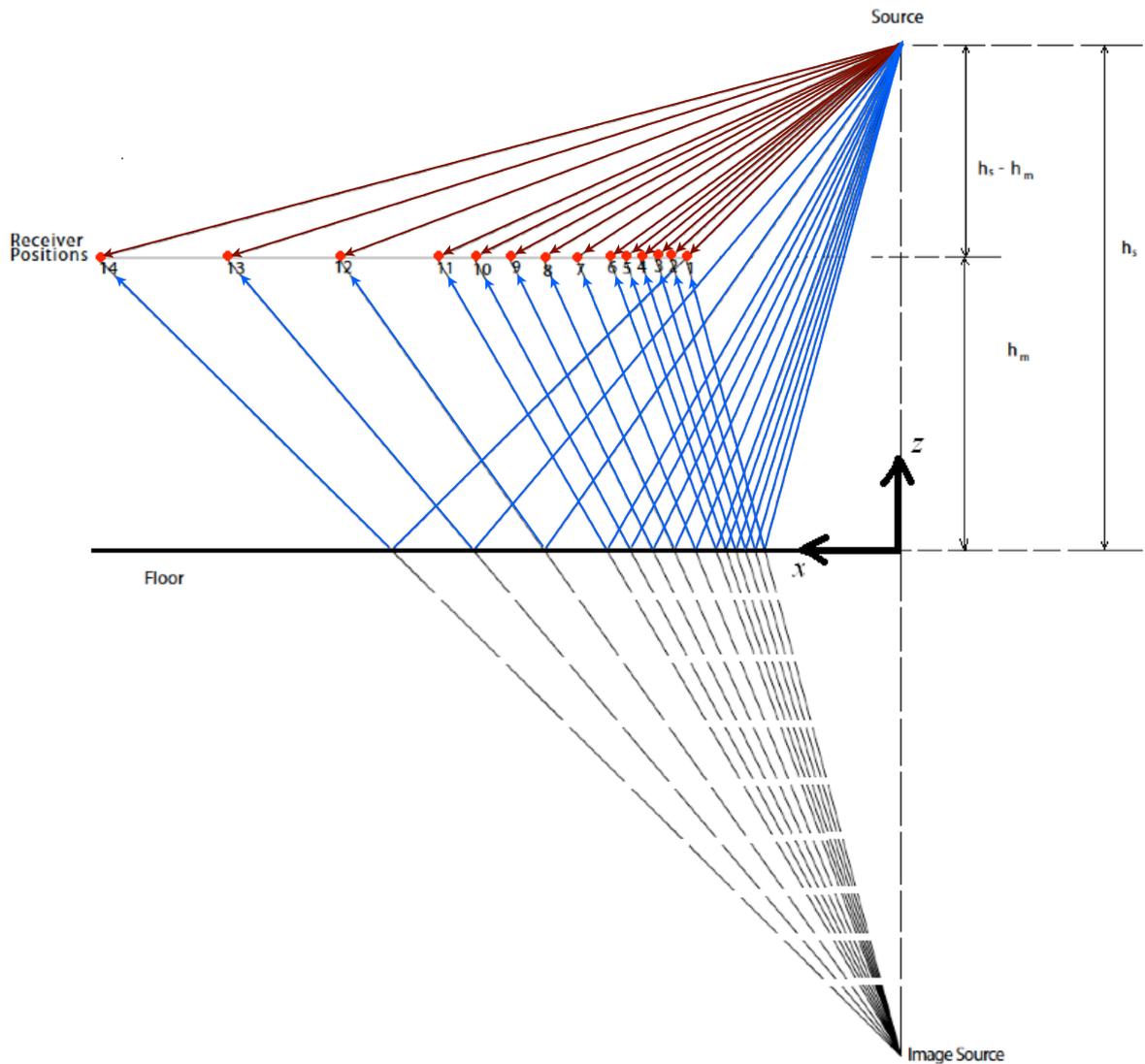
Figure 5.9: Image source model used for determining the angle of arrival of first reflections (θ_{m_i}).

The image source model is useful to visualise the *TOF* of direct sound and first reflection. For the positions, which are far from the source e.g. $P_{11} - P_{14}$, these distances are approaching to be similar. This has some implications on the distribution of the impulse response. To familiarise with the data used for calculation, please refer to the values of the angles of arrival of direct sound and first reflection to **Table 5.1** for the coordinates in *x*-axis and the associated arrival angle of direct sound to **Table 5.2** and for the associated arrival angle of first reflection to **Table 5.3**.
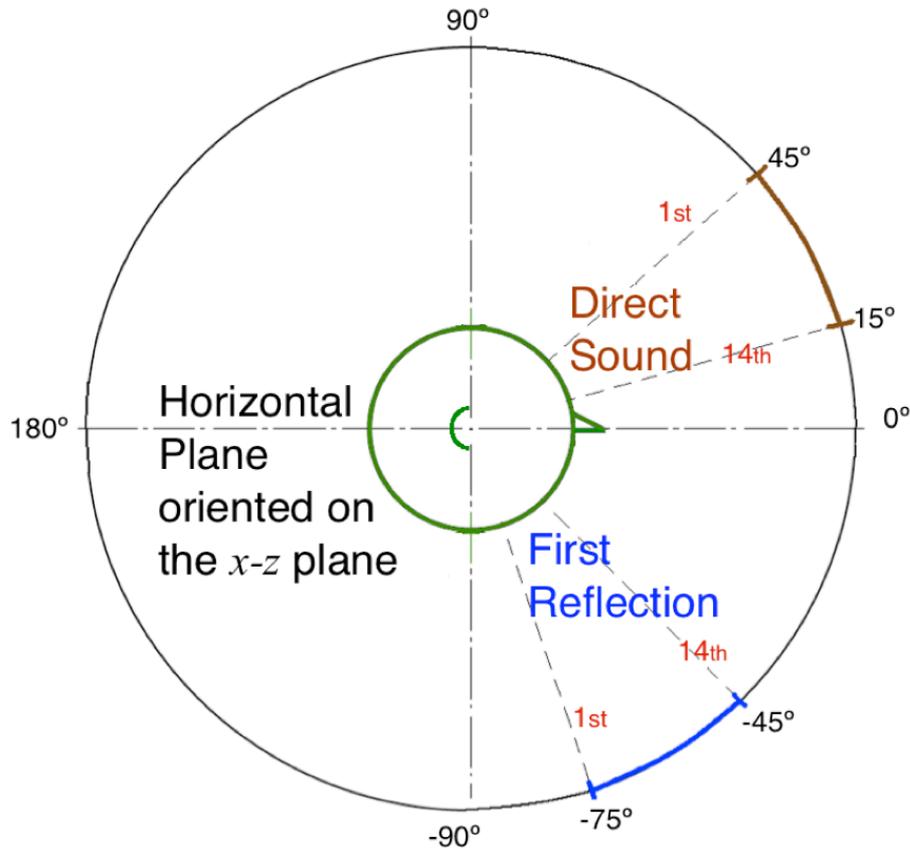
Figure 5.10: Range of angles of arrival of direct sound (θ_{s_i}) and first reflection (θ_{m_i}) tested in *x-z* (vertical plane).

5.2.1.1 *Soundfield* microphone model ST 350 used to measure single reflection scenario

This test uses the *Soundfield* microphone model ST 350 measuring 14 positions ($P_1 - P_{14}$). Each position was repeated 10 times in order to compute statistics of repeatability. For each position, an estimation of the angle of arrival of direct sound (θ_{s_i}) and first reflection (θ_{m_i}) was performed by two main methods: a) time domain method (*TD*) and b) short-time Fourier transform (*STFT*) method.

5.2.1.2 One-dimensional face-to-face p-p intensity probe rotated

This experiment uses the face-to-face p-p probe measuring the same 14 positions depicted in **Figure 5.8**. Each position was repeated 10 times in order to compute statistics of repeatability. For each position, an estimation of the angle of arrival of direct sound (θ_{s_i}) and first reflection (θ_{m_i}) was performed by two main methods: a) time domain method (TD) and b) short-time Fourier transform ($STFT$) method.

The spacing of p-p intensity probe directly affects the range of frequency. In order to investigate which spacing was the optimum for the specifications of the needed frequency range, some variations were proposed and measured. Three distances were tested in practical measurements: a) 10.3 mm with distance spacer, b) 10.3 mm without distance spacer, c) 50 mm without distance spacer, and d) 3.10 mm without distance spacer. If the analysis of the first position was not localized properly with time domain and short-time Fourier transform ($STFT$) methods, no more analysis of positions are performed. Among all the positions tested, the results showed that only the configuration of the 10.3 mm solid spacer was capable of delivering accurate results within a frequency range of approximately 125 Hz up to 5 kHz (Gade, 1982b, Gade, 1982a). The following photos are some of the configurations used for measuring with the p-p intensity probe:

Figure 5.11: Different configurations of spacers tested on the p-p probe.

    a) $d = 10.33$ mm   b) $d = 10.33$ mm   c) $d = 50.00$ mm   d) $d = 3.10$ mm

    2 microphones   1 single microphone.   2 microphones   2 microphones.

According to Watkinson and Fahy (Watkinson and Fahy, 1984) and later validated by Jacobsen et al. (Jacobsen et al., 1996), a spacer of 12 mm seems to be a good compromise between 50 Hz up to 10 kHz as a wide band intensity p-p intensity probe. When the diameter of the spacer is equal to the distance of the spacer, there is an extra octave, which is attained in the high frequency limit. Nevertheless, in practice the p-p intensity probe used for the measurements showed inadequate performance below 700 Hz.

5.3 Single microphone multiple positions

Unfortunately, this case was the least accurate of the four configurations of p-p intensity probe's spacers tested. Results are not shown since the error found for the direct sound was greater than the tolerance of minimum audible angle (MAA). It needs to be very precise in the setup of each of the four measurements and ensure to perform the measurements with care of time invariance. (See **Figure 5.11** on section b) and **Figure 5.12**.)
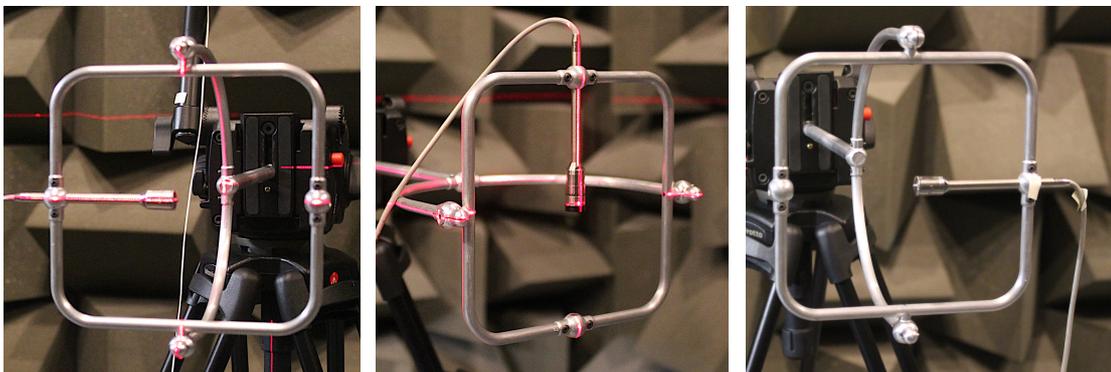
Figure 5.12: Example of a set of 3 from 4 rotations in clockwise direction needed to obtain a single 2-D intensity at a single point in space.

5.4 Multiple reflections scenario in laboratory conditions

Simultaneously arriving reflections may occur when two sound waves arrive at the same discrete time index (n) from different angles (θ). The resultant magnitude and direction of concurrent reflections will be incorrect, since the procedure typically employed to add

the components on each axis produces a single resultant vector. Therefore, it is impossible to differentiate simultaneous reflections in time domain (TD). To create this measurement scenario, two reflector panels are located parallel to each other, additionally they are parallel to the line that joins the source and the receiver in the semi-anechoic chamber. This measurement scenario was investigated with practical measurements by Abdou (Abdou, 1994) using a one-dimensional p-p intensity probe rotated in orthogonal axis. Later, it was simulated by a mathematical model and later physically measured by Gover with a 32-microphone array using beamforming techniques (Gover, 2001). The conclusions are that no matter how accurate the coincident microphone array is or the intensity probe, the measurement system is not fully capable of resolving true simultaneous reflections in the time domain (TD). The reason for this is because under these conditions, the reflections will fuse together and create an aberrant new vector that is not taking into account the real directions of the previous simultaneous vectors. The instantaneous intensity ($\vec{i}_{inst}$) will only be useful to describe the direction and strength of the net transport of energy (Burns, 1996), but strictly speaking, it will not be able to differentiate true simultaneous arrivals. In that case, the only way to overcome this situation is by using beamforming and cross-correlation between the microphones of a spaced array (Tervo, 2012, Protheroe et al., 2013).

The simultaneous arrival of reflections was studied by Van Lancker in (van Lancker, 2000) using a microphone array consisting of 8 omni-directional electret microphones located at each apex of a cube, with a distance of 25 cm between each microphone. The location of the direct sound and the early reflections was performed using time-delay estimation (TDE) techniques. In spite of the fact that he only reports the accuracy of the measuring system on the direct sound with a DOA down to 1 or 2 degrees, he does not report the early reflection DOA estimation error.

Nevertheless, the computational cost of using beamforming makes a real time measurement system difficult to implement. For the consumer's market it is also an expensive solution. The methodology used for this project focuses on overcoming this issue in a practical by way by increasing the sample rate of the analogue to digital converter to the highest resolution available (currently 192 kHz) and apply the short-time

Fourier transform method to the signal in order to investigate differences on the spectrum of the arrival signals. This may provide differences that will enable the resolution of some of the simultaneous reflections. This potentially increases the accuracy of the estimation of angle, and is likely to be able to resolve most of the crucial simultaneous reflections, which appear in the early reflection stage of the decay. It is important to note that in this thesis, the case of simultaneous reflections was not implemented in practice.

5.5 Three-dimensional impulse response (3-D *IR*) measurements in small rooms

A practical use of the acquisition of 3-D impulse responses (3-D *IR*) is to measure the distribution of reflections during the decay time of any room and map it to describe it in different planes. Since this thesis is focusing in small rooms, the first target was to measure a control room and analyse its temporal decay. Specifically, how the room evolves from a very directive sound field when the direct sound arrives to the receiver until the sound in the room vanishes. One of the desirable properties of a control room is to be able to hear the direct sound and avoid flutter echoes and strong early reflections. By applying the analysis of the impulse response (*IR*) using the *STFT* method is possible to analyse the diffuseness and assess the temporal and spectral properties of the sound decay. This experiment was done before having access to the p-p probe, therefore only the *Soundfield* was used. A description of the measurement conditions follows:

A sample of 10 measurements of a 3-D *IR* in small rooms was done using the *Soundfield* model SPS 422B microphone and applying the short Fourier transform method (*STFT*). **Figure 5.13** shows the measurement scenario.

Only the horizontal plane (*xy*-axis) is investigated at this time. The measurements were made in Blue Room 2 control Room from the facilities of recording studios at the University of Huddersfield, whose dimensions are $4.16 \times 2.75 \times 2.87$ m and its volume is 32.1 m^3. Measurements were performed at two different source positions. The first was at the front of the microphone and the second at 120º from the origin, referenced at front with an approximated angle of 0º (which was later measured and was 2º, see **Figure**

4.12, **Figure 7.1** and **Figure 7.2**). The sampling frequency resolution was set to 96 kHz, which gives a temporal resolution of 0.01 ms.

This room has a fairly short reverberation time RT = 0.17 s @ 1 kHz, and a Mean Free Path l_{MFP} = 2.2 m giving a minimum time window of Δt = 6 ms. However, smaller time windows of 0.1 and 1 ms were tested in this analysis and seem more appropriate to determine any isolated early reflection. Application of the *Soundfield* microphone model SP 422 to small rooms. The results of this experiment are covered in Chapter 7 and the discussion of these results is found in Chapter 8,(section 8.2.1).
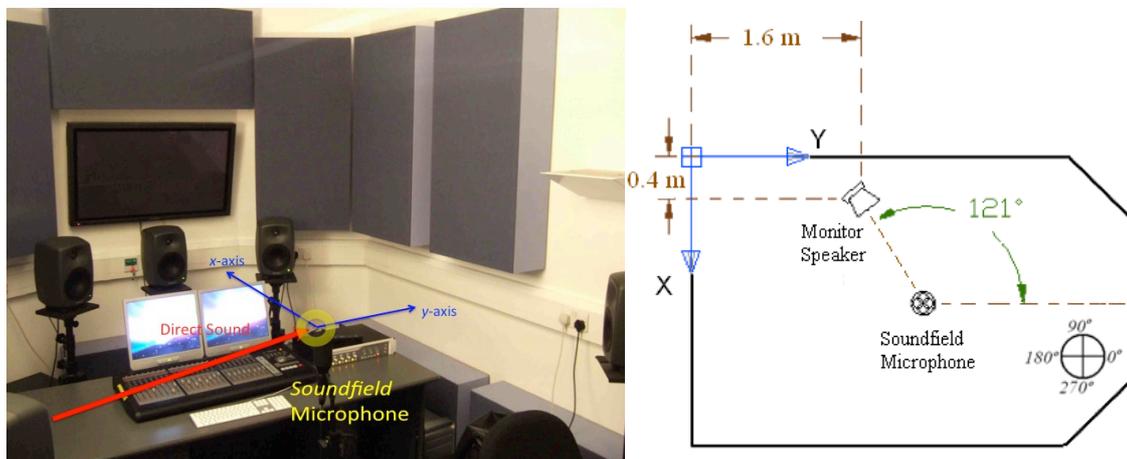
Figure 5.13: Photo and sketch of Blue room 2 measured with the *Soundfield* microphone model SP 422.

5.6 Summary

This section describes details of the Acoustic probe configurations, covering the measurement conditions and the repetitions needed to validate the accuracy of the method. The next section covered is the description of the measurement environment, which is comprised of: semi-anechoic chamber, monitor speaker influence, damping vibration on source and receiver, and the laser cross system to define spatial coordinates. It describes the laboratory test for single reflection measurements in the semi-anechoic chamber using the p-p probe and the *Soundfield* microphone. Afterwards, the single microphone multiple positions is mentioned, which was discarded because of bad

performance in angular accuracy of sound detection. An explanation of the effect of varying the spacing between the microphones is depicted in that section. The usable frequency range is discussed according to the optimum selection of spacing at 10.3 mm. The next topic covered is the applications of 3-D *IR* measurements in a control room where the application of a cheaper and more versatile *Soundfield* microphone for the measurement of a real room response with multiple reflections.

Chapter 6 : Results of laboratory tests

This section presents the results of the experiments undertaken in the single reflection scenario (for an explanation of it please refer to section 5.2.1). The case of measuring a single reflection in the semi-anechoic chamber is tested with the time domain (TD) method and the frequency domain method ($STFT$).

The detection of the angle of arrival of direct sound (θ_{s_i}) is a trivial issue, but the detection of the angle of arrival of first reflection (θ_{m_i}) is not, and usually exhibits larger errors in its estimation. Therefore, it is expected to have a larger errors and variations of the positions, which have the longest distance paths because the time of flight of the direct sound and the first reflection is closer than when the path difference is longer (Defrance et al., 2009). Owing to the fact of the high sensitivity of the system to the angle of arrival of first reflection and its tendency to deliver errors on the estimation, two measurement methods were used. This is because the time domain method (TD) and the frequency domain method ($STFT$) complement each other's deficiencies.

Due to the great amount of data created in this experiment, only the mean absolute percentage error graphs and their uncertainty are presented. Initially the most intuitive graphs were the error graphs. They showed if the measure was over estimating the real value or if it was under estimating. Later the error graphs showed an analysis of the errors with positive or negative sign. This complicated the analysis of errors. It was needed to use an absolute norm of error. Mean squared error was discharged because it introduced a non-linear explanation of errors while mean absolute percentage error seem to preserve the error without introducing any non-linear behaviour.
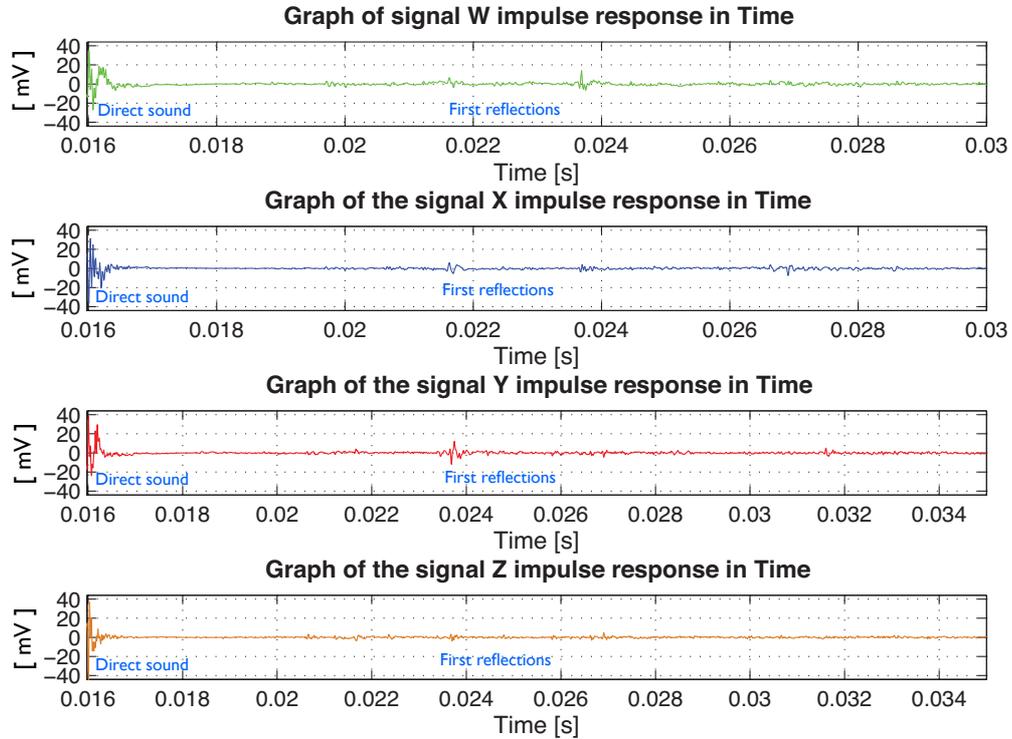
Figure 6.1: Example of impulse response B-Format signals measured. It contains the direct sound and first reflection (signals not normalised).

6.1 Accuracy of methods of analysis

6.1.1 The measurements using the *Soundfield* microphone model ST 350

This procedure uses the *Soundfield* microphone model ST 350 and compute circular statistics for repeatability and determination of the uncertainty of measurements by using the circular mean ($\bar{\theta}$) and the circular standard deviation (σ_o) and the computation of the *MAPE* and weighted average of standards deviation of the *MAPE* ($\bar{\sigma}_{\bar{\varepsilon}_\%}$). For each position, an estimation of the angle of arrival of direct sound (θ_{s_i}) and first reflection (θ_{m_i}) was performed by two main methods: a) time domain method (*TD*) and b) short-time

Fourier transform (*STFT*) method, for more details refer to **Figure 6.2**, **Figure 6.3** and **Figure 6.4**.
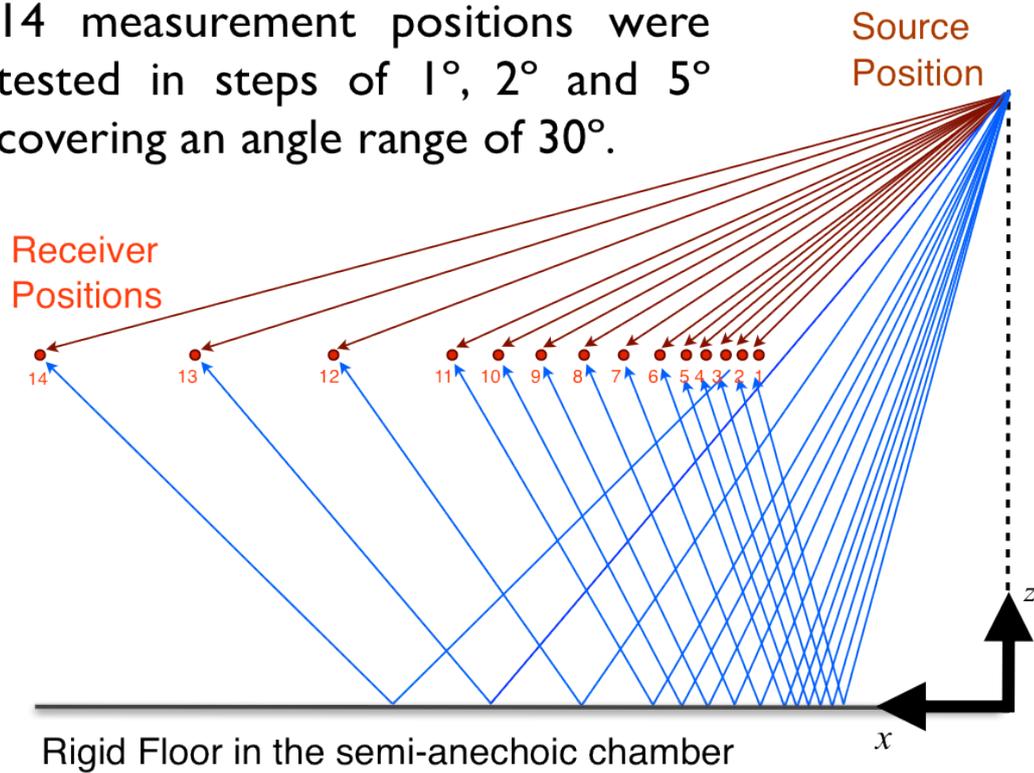
Figure 6.2: Sketch of the 14 positions ($P_1 - P_{14}$) for reference of the results.
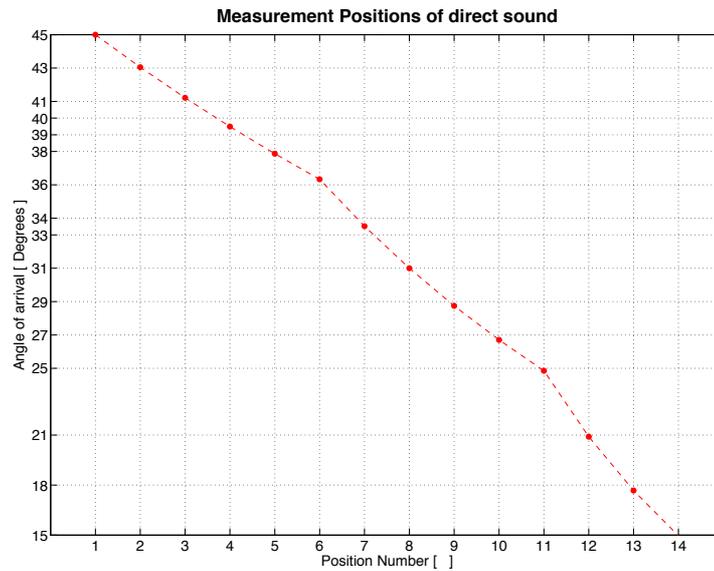
Figure 6.3: Graph of the measurement positions results for the direct sound. According to the steps of 1º for position $P_1 - P_5$, then 2º for position $P_6 - P_{11}$ and finally 5º for position $P_{11} - P_{14}$.
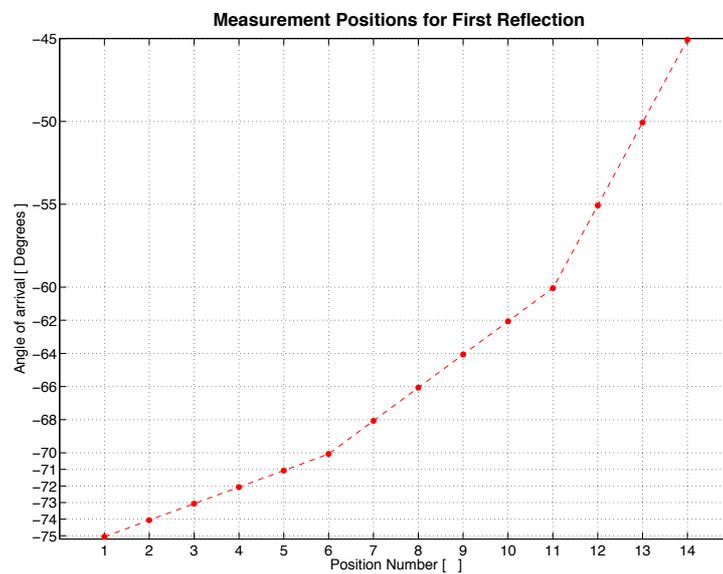
Figure 6.4: Graph of the measurement positions for the first reflection. According to the steps of 1º for position $P_1 - P_5$, then 2º for position $P_6 - P_{11}$ and finally 5º for position $P_{11} - P_{14}$.

6.1.1.1 Time domain method using instantaneous intensity ($\vec{i}_{inst}$) using the *Soundfield* microphone
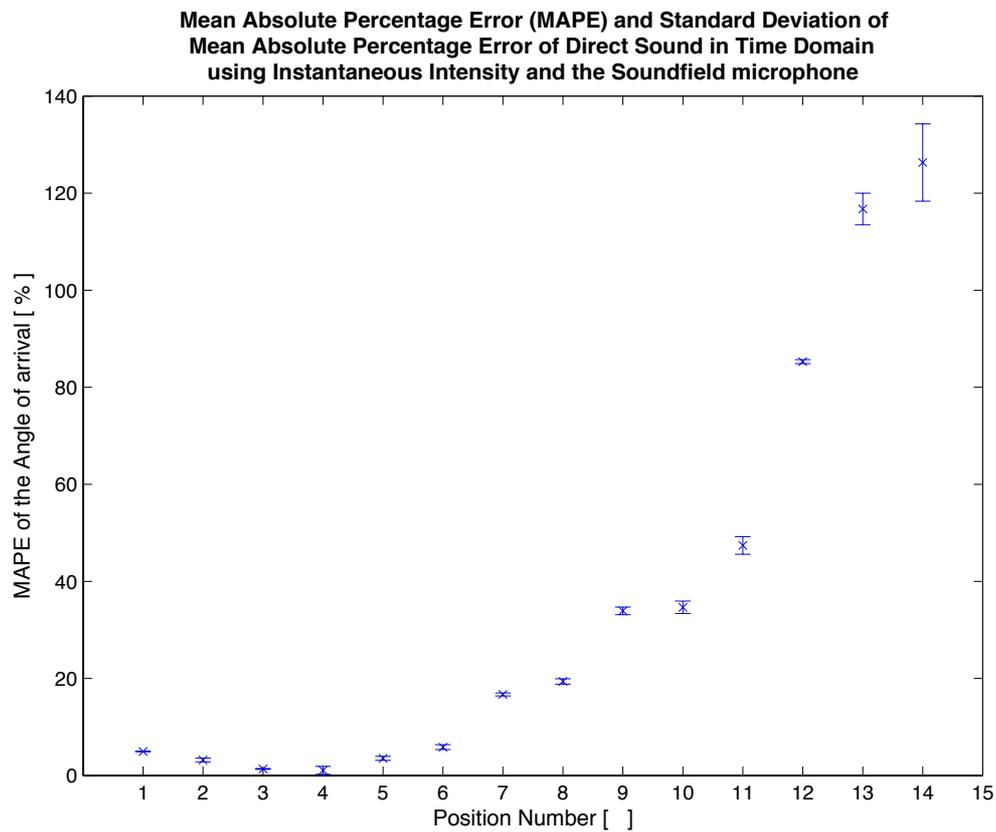
Figure 6.5: Mean absolute percentage error ($MAPE$) $\left|\bar{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) of direct sound using instantaneous intensity ($\vec{i}_{inst}$) in time domain (TD) using the *Soundfield* microphone model ST 350.
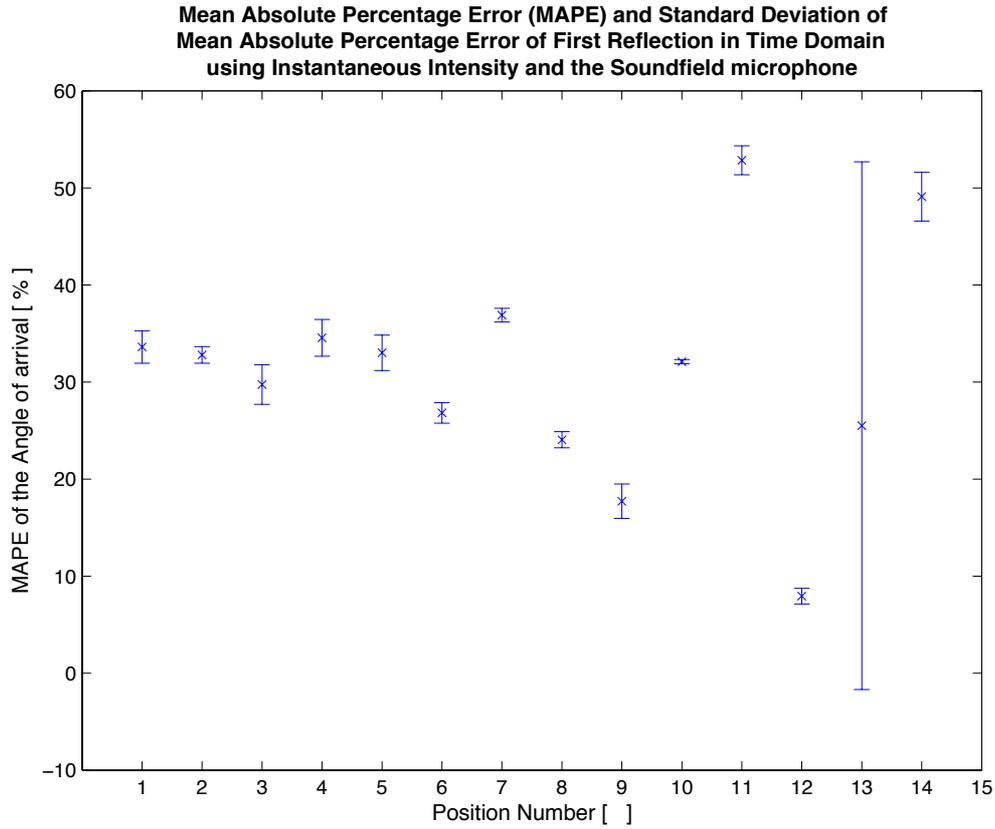
Figure 6.6: Mean absolute percentage error ($MAPE$) $\left|\bar{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) of first reflection using instantaneous intensity ($\vec{i}_{inst}$) in time domain (TD) using the *Soundfield* microphone model ST 350.

Generally, inaccurate results were obtained with positions P_6 to P_{12} in direct sound and first reflection. In specific position P_{13} and P_{14}, reported errors are somehow too high to be used. These last positions exhibit shallow first reflection arrival angles (θ_{m_i}) that tend to be difficult to measure by the microphone. One of the reasons could be insufficient signal-to-noise-ratio (SNR), the other is a smearing of the synchronicity between channels that creates estimation errors.

1) The direct sound overall absolute mean error $\left|\bar{\varepsilon}_{\bar{\varepsilon}}\right|$ for the direct sound is 7.071º, which corresponds to an overall mean absolute percentage error $\left|\bar{\varepsilon}_{\bar{\varepsilon}_{\%}}\right|$ of 35.724% and an overall weighted average of standard deviation ($\bar{\sigma}_{\bar{\varepsilon}}$) of ±0.207º, which corresponds to an overall weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) of ±1.833%.

2) The overall absolute mean error $\left|\bar{\varepsilon}_{\bar{\varepsilon}}\right|$ for the first reflection is 19.113º, which corresponds to an overall mean absolute percentage error $\left|\bar{\varepsilon}_{\bar{\varepsilon}_{\%}}\right|$ of 31.185% and an overall weighted average of standard deviation ($\bar{\sigma}_{\bar{\varepsilon}}$) of ±2.416º, which corresponds to an overall weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) of ±3.629%.

6.1.1.2 Time domain method using the complex instantaneous intensity ($\vec{i}_{complex}$) using the *Soundfield* microphone
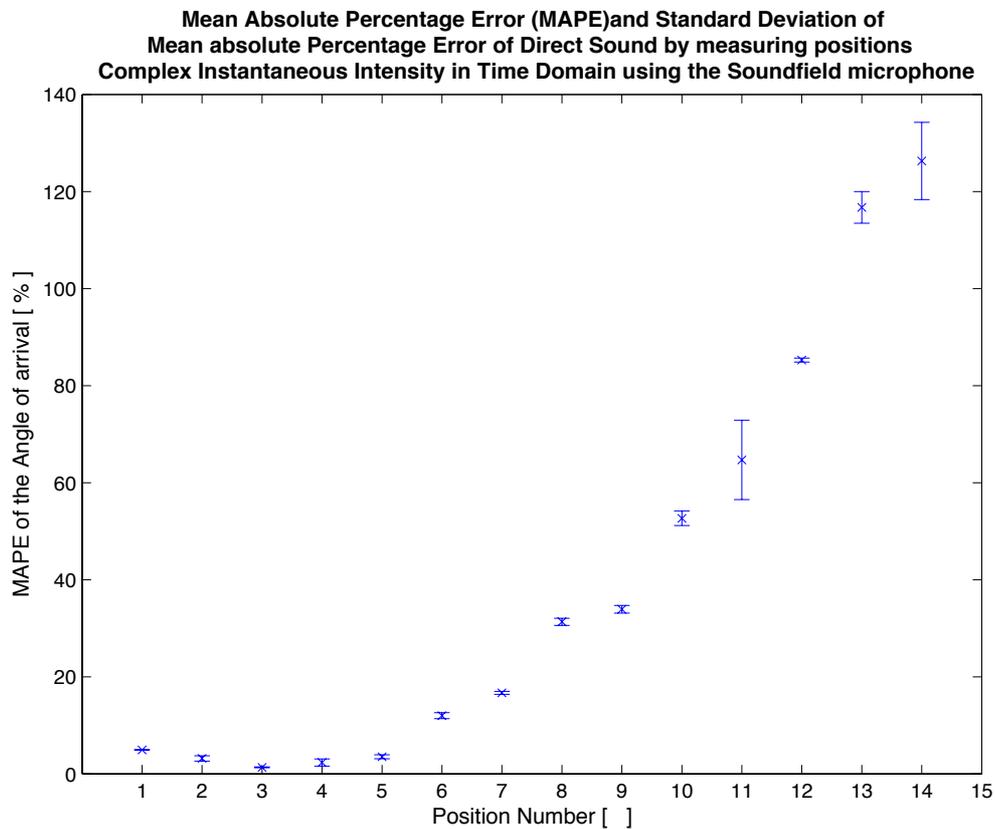
Figure 6.7: Mean absolute percentage error ($MAPE$) $\left| \overline{\varepsilon}_{\%} \right|$ and weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) of direct sound using complex instantaneous intensity ($\vec{i}_{complex}$) in time domain (TD) using the *Soundfield* microphone model ST 350.
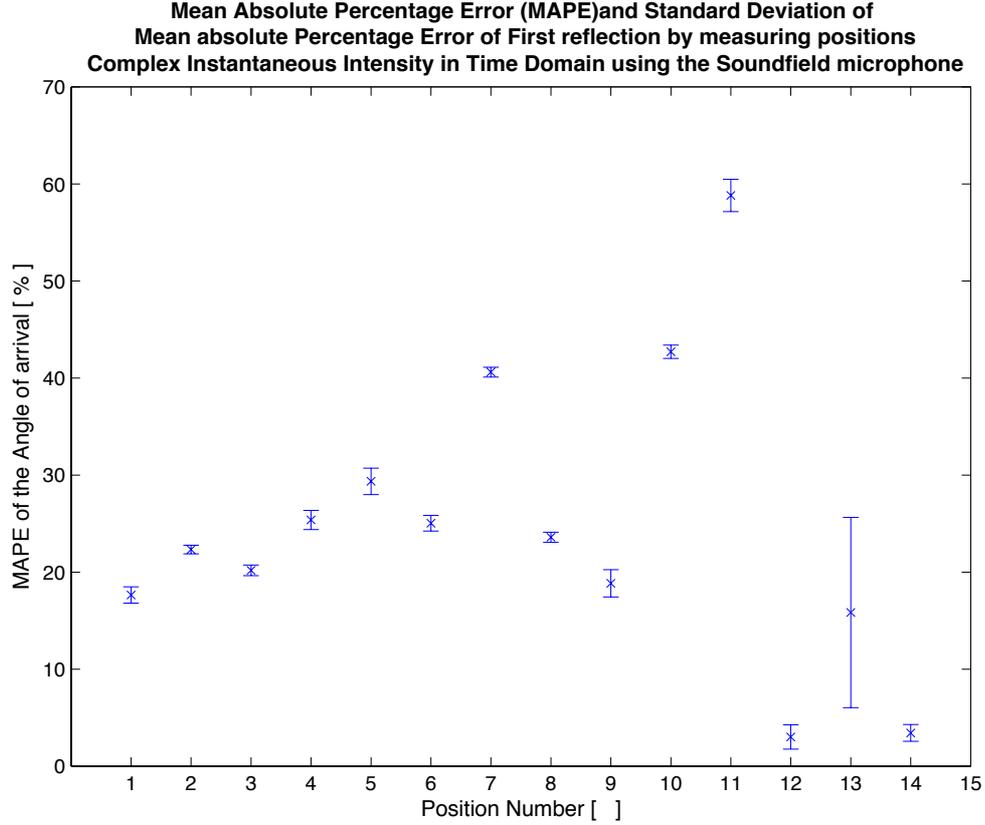
Figure 6.8: Mean Absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) of first reflection using complex instantaneous intensity ($\vec{i}_{complex}$) using the *Soundfield* microphone model ST 350.

1) The direct sound overall absolute mean error $\left|\overline{\varepsilon}_{\overline{\varepsilon}}\right|$ for the direct sound is 8.208º, which corresponds to an absolute mean percentage error $\left|\overline{\varepsilon}_{\%}\right|$ of 39.630% and a weighted average of standard deviation error ($\overline{\sigma}_{\overline{\varepsilon}}$) of ±0.330º, which corresponds to an overall weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) of ±1.580%.

2) The overall absolute mean error $\left|\overline{\varepsilon}_{\overline{\varepsilon}}\right|$ for the first reflection is 15.895º, which corresponds to a mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ of 24.764% with a weighted average of standard deviation error ($\overline{\sigma}_{\varepsilon}$) of ±1.248º, which corresponds to a

weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) of $\pm 1.366\%$.

6.1.1.3 Short-time Fourier transform ($STFT$) method using active instantaneous intensity ($\vec{I}_a$) using the *Soundfield* microphone

The following graphs are the mean absolute percentage errors ($MAPE$) and weighted average of standard deviation of mean absolute percentage errors for direct sound and first reflections across the 16 ERB bands using the *Soundfield* microphone model ST350.
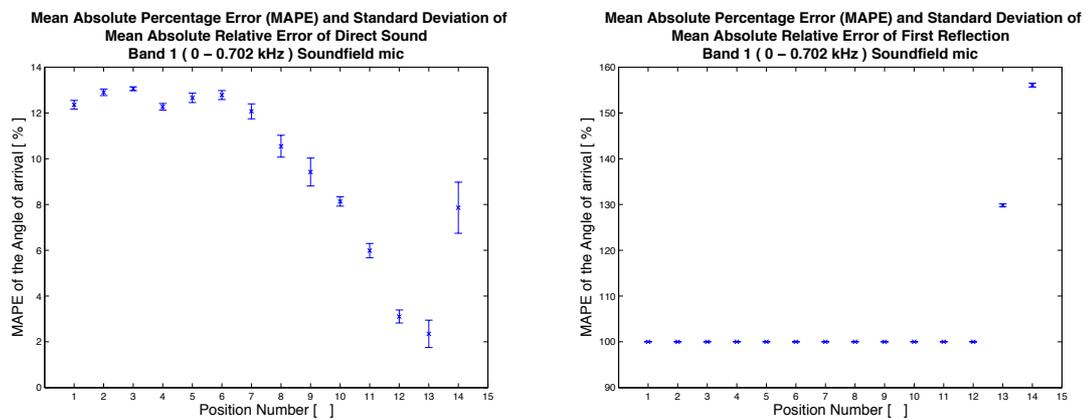
Figure 6.9: *ERB* band 1 mean absolute percentage errors $\left|\bar{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
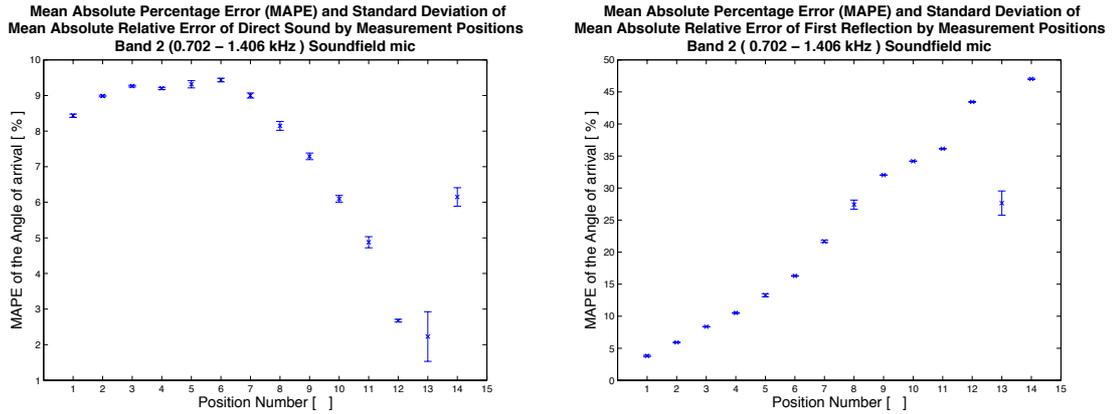
Figure 6.10: *ERB* band 2 mean absolute percentage errors $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
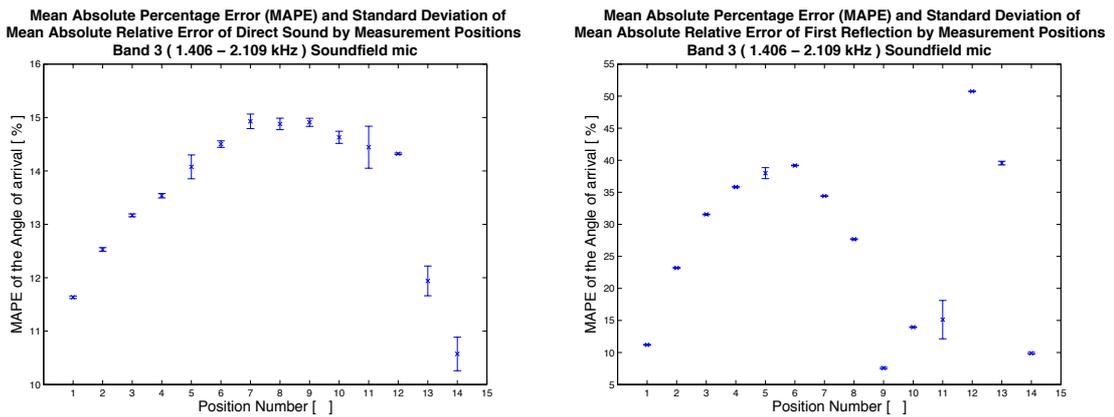
Figure 6.11: *ERB* band 3 mean absolute percentage errors $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
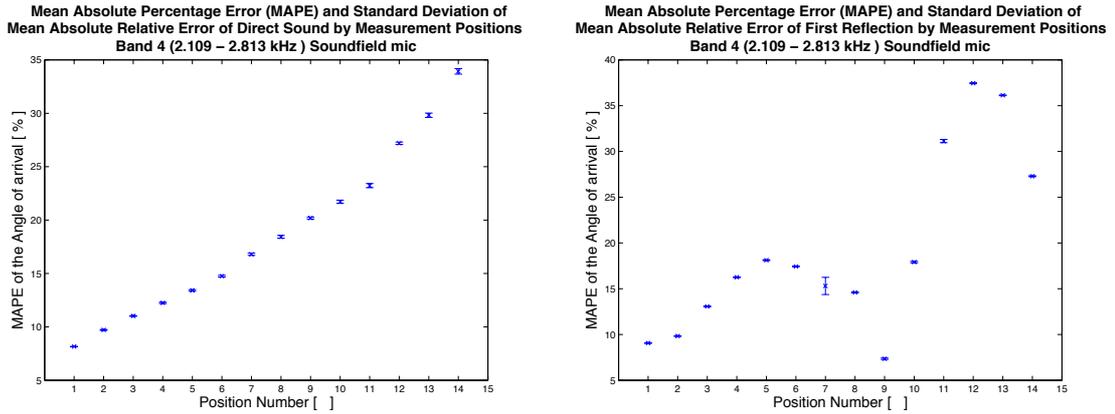
Figure 6.12: *ERB* band 4 mean absolute percentage errors $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
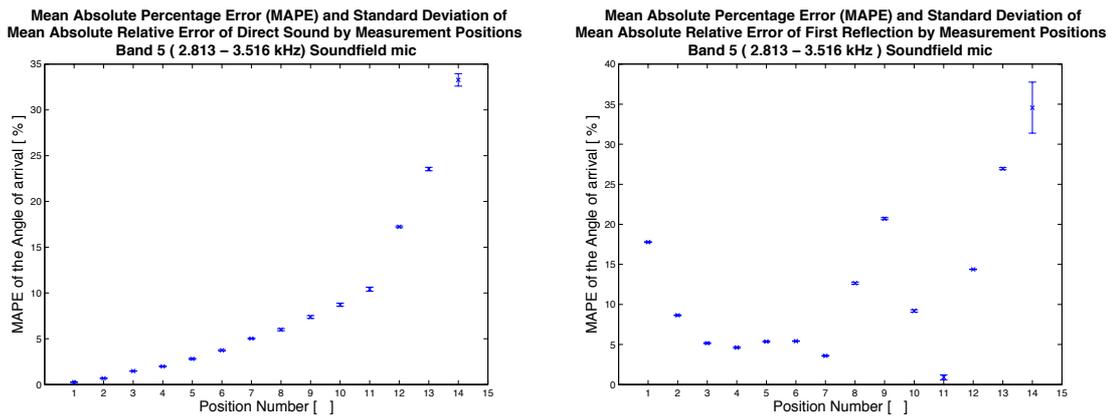
Figure 6.13: *ERB* band 5 mean absolute percentage errors $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
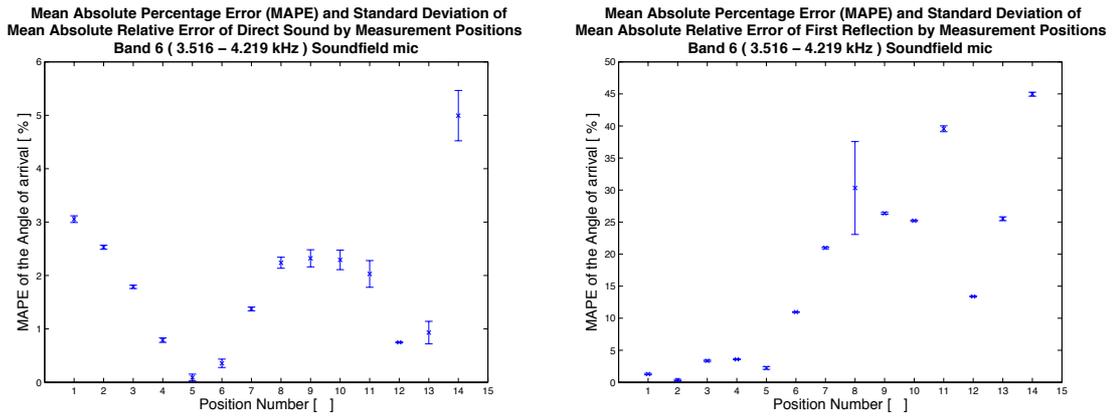
Figure 6.14: *ERB* band 6 mean absolute percentage errors $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
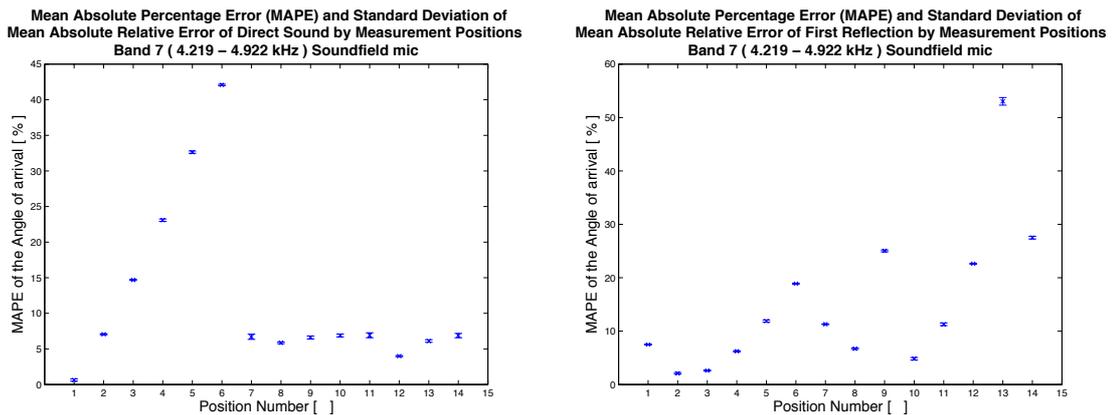
Figure 6.15: *ERB* band 7 mean absolute percentage errors $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
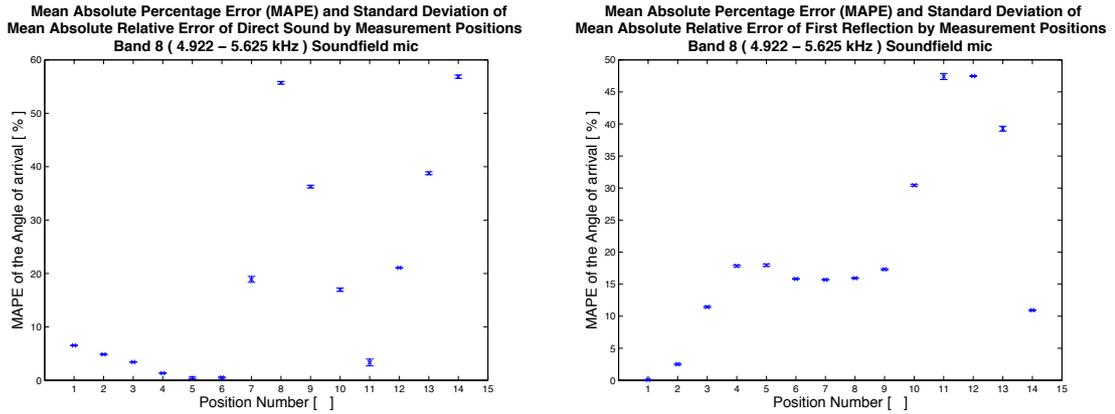
Figure 6.16: *ERB* band 8 mean absolute percentage errors $\left|\bar{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
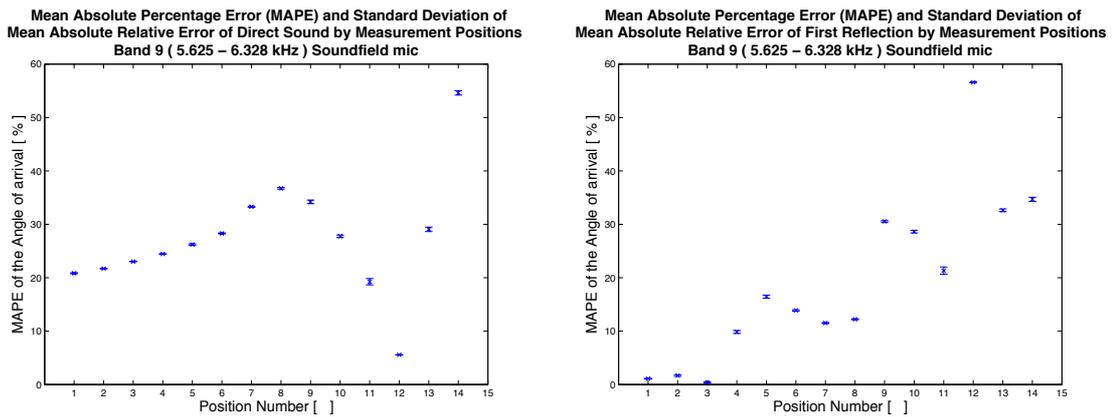
Figure 6.17: *ERB* band 9 mean absolute percentage errors $\left|\bar{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
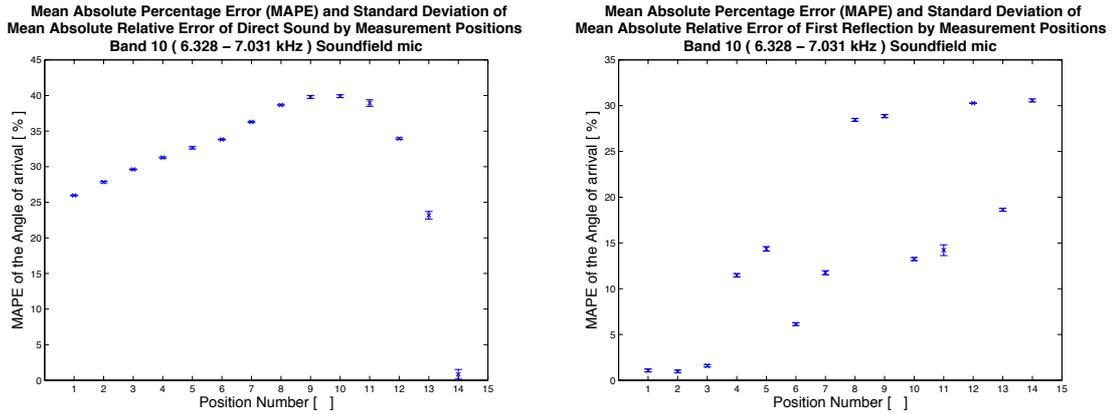
Figure 6.18: *ERB* band 10 mean absolute percentage errors $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
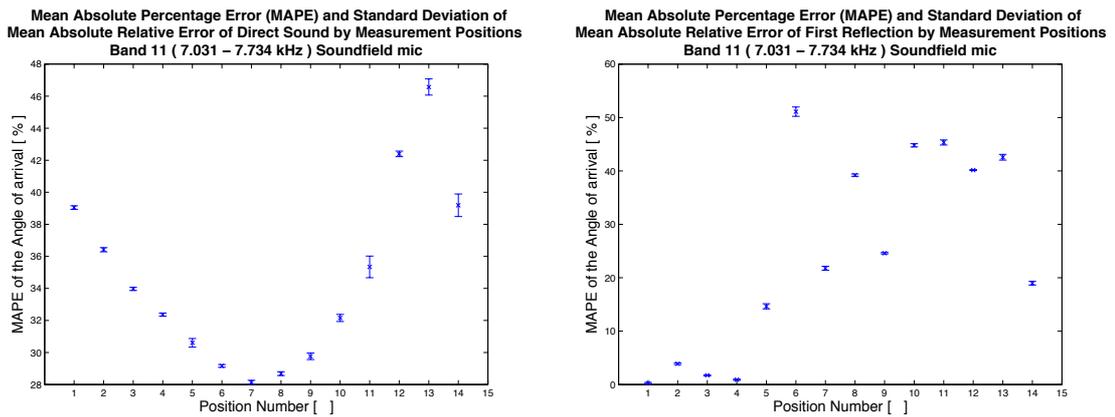
Figure 6.19: *ERB* band 11 mean absolute percentage errors $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
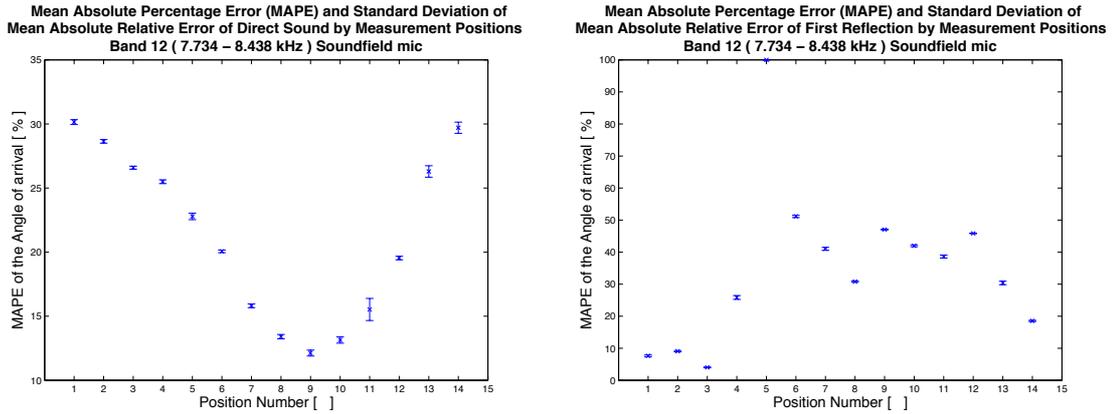
Figure 6.20: *ERB* band 12 mean absolute percentage errors $\left|\bar{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
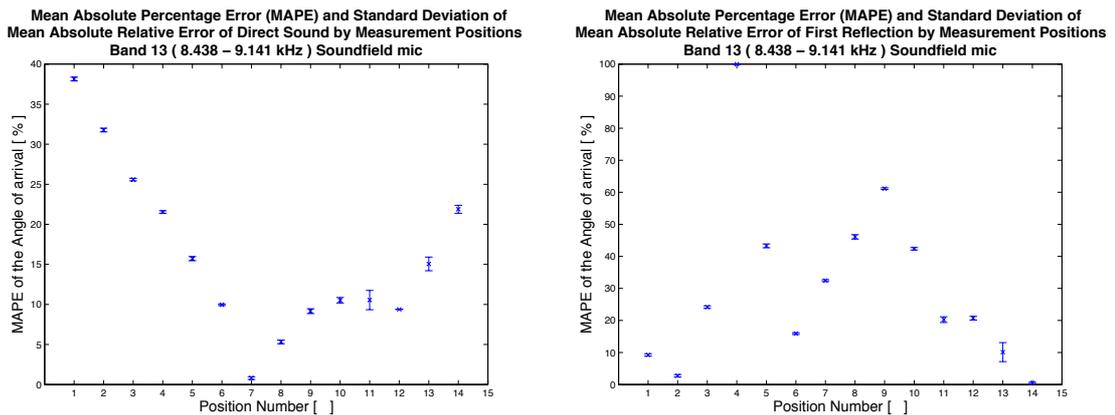
Figure 6.21: *ERB* band 13 mean absolute percentage errors $\left|\bar{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
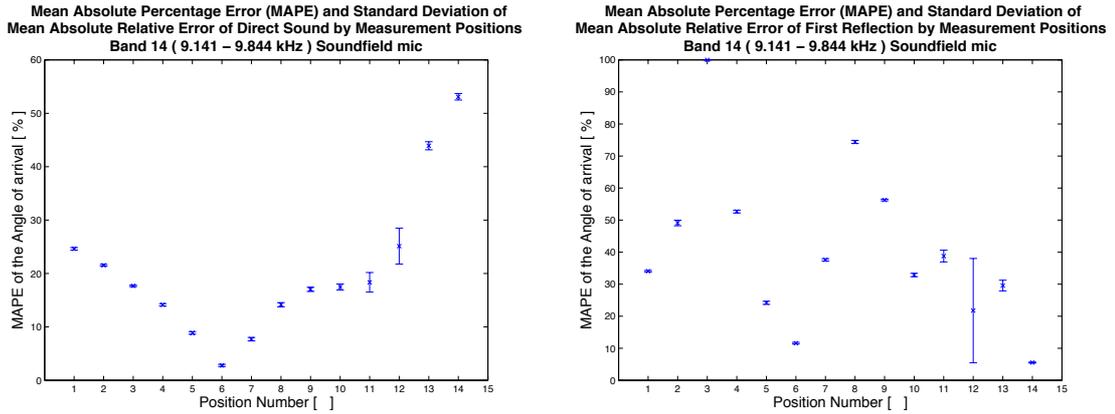
Figure 6.22: *ERB* band 14 mean absolute percentage errors $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
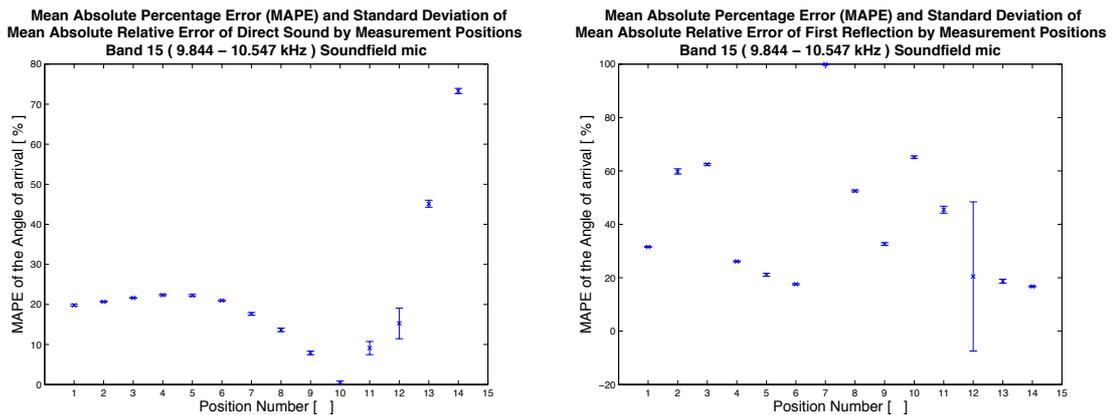
Figure 6.23: *ERB* band 15 mean absolute percentage errors $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average of standard deviation of mean absolute percentage errors ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.
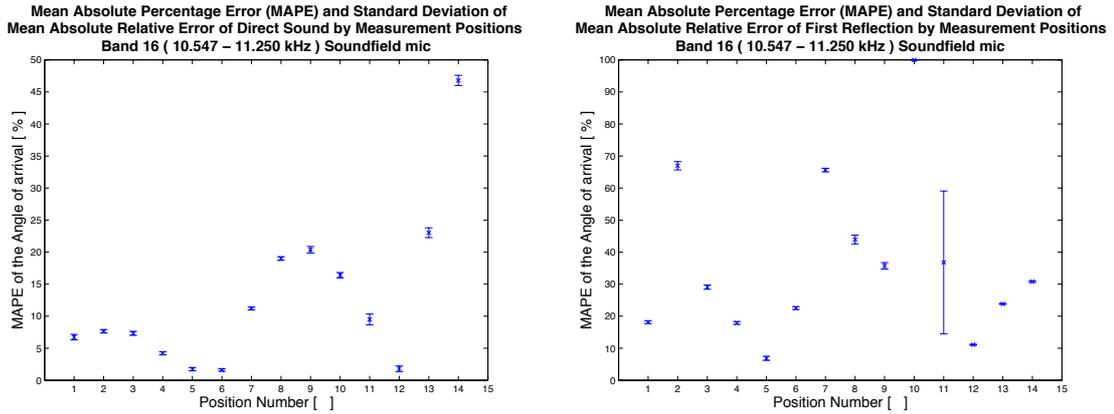
Figure 6.24: *ERB* band 16 mean absolute percentage errors $\left| \bar{\varepsilon}_{\%} \right|$ and weighted average of standard deviation of mean absolute percentage errors ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) for direct sound (left) and first reflection (right) for *Soundfield* microphone.

The main results are summarised in the following data:

1) The direct sound overall mean error ($\bar{\varepsilon}_{\bar{\varepsilon}}$) for the direct sound is 1.351º, which in terms of absolute mean absolute percentage error $\left| \bar{\varepsilon}_{\%} \right|$ is 17.280% and a weighted average of standard deviation error ($\bar{\sigma}_{\bar{\varepsilon}}$) of ±0.013º, which in terms of weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) represents ±0.060%.

2) The overall mean error ($\bar{\varepsilon}_{\bar{\varepsilon}}$) for the first reflection is 18.346º, which in terms of absolute mean absolute percentage error $\left| \bar{\varepsilon}_{\%} \right|$ is 30.286% with a weighted average of standard deviation error ($\bar{\sigma}_{\bar{\varepsilon}}$) of ±0.198º, which in terms of weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) represents ±1.333%.

After the analysis of the performance of the *Soundfield* microphone by 16 *ERB* bands using the direct sound and the first reflection, it was found that a synthesis of these data in a graph can show the bands that perform better, and it is possible to select the frequency working range of the system:
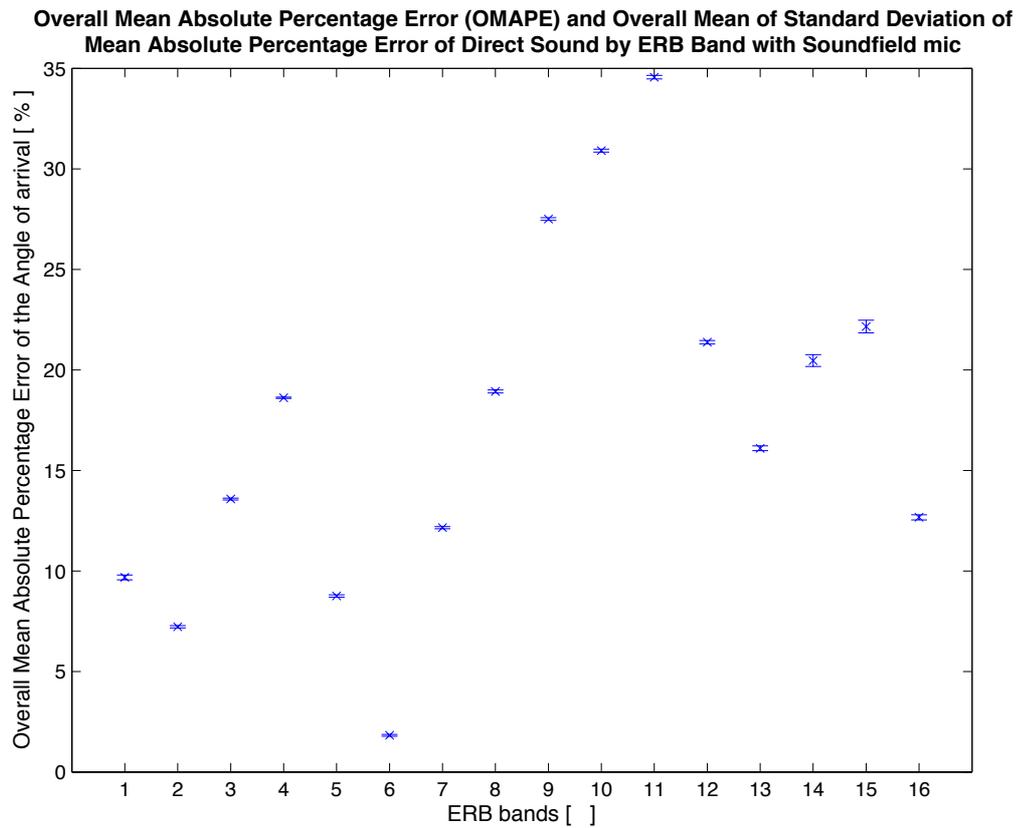
Figure 6.25: Overall mean absolute percentage error ($OMAPE$) $\left|\overline{\varepsilon}_{\overline{\varepsilon}_{\%}}\right|$ and overall weighted average of standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ of direct sound by *ERB* band for the direct sound.
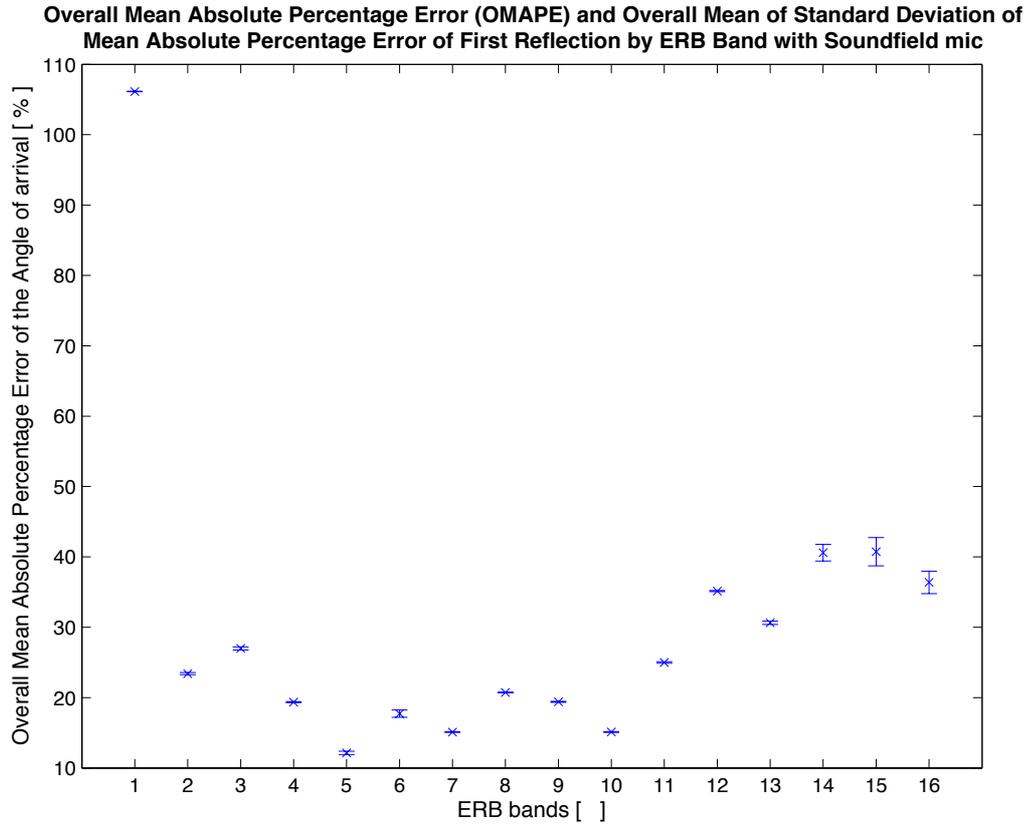
Figure 6.26: Overall mean absolute percentage error ($OMAPE$) $\left|\bar{\varepsilon}_{\bar{\varepsilon}_{\%}}\right|$ and overall weighted average of standard deviation of mean absolute percentage error $\bar{\sigma}_{\bar{\varepsilon}_{\%}}$ of direct sound by ERB band for the first reflection.

6.1.1.4 Average of each position using short-time Fourier transform ($STFT$) using the *Soundfield* microphone

Since poor results were obtained using both prior methods, a new descriptor was studied. It consists of extracting the angle of the arrival from the *STFT* method and performing an average of all the *ERB* bands at the data of direct sound and first reflection to obtain a single mean estimation angle per each receiver position. It was more accurate than the other two in the estimation of the angle on the position that were close to the source. However, in the case of the direct sound the error is relatively small, and overestimates, and underestimates the incoming angle depending on the distance from the source. In the case of the first reflection, the descriptor overestimated the reference values.

Overall Mean Absolute Percentage Error (OMAPE) and Average Standard Deviation of Mean Absolute Percentage Error of Direct Sound by Measurement Positions Average Instantaneous Intensity by each Position using the Soundfield microphone
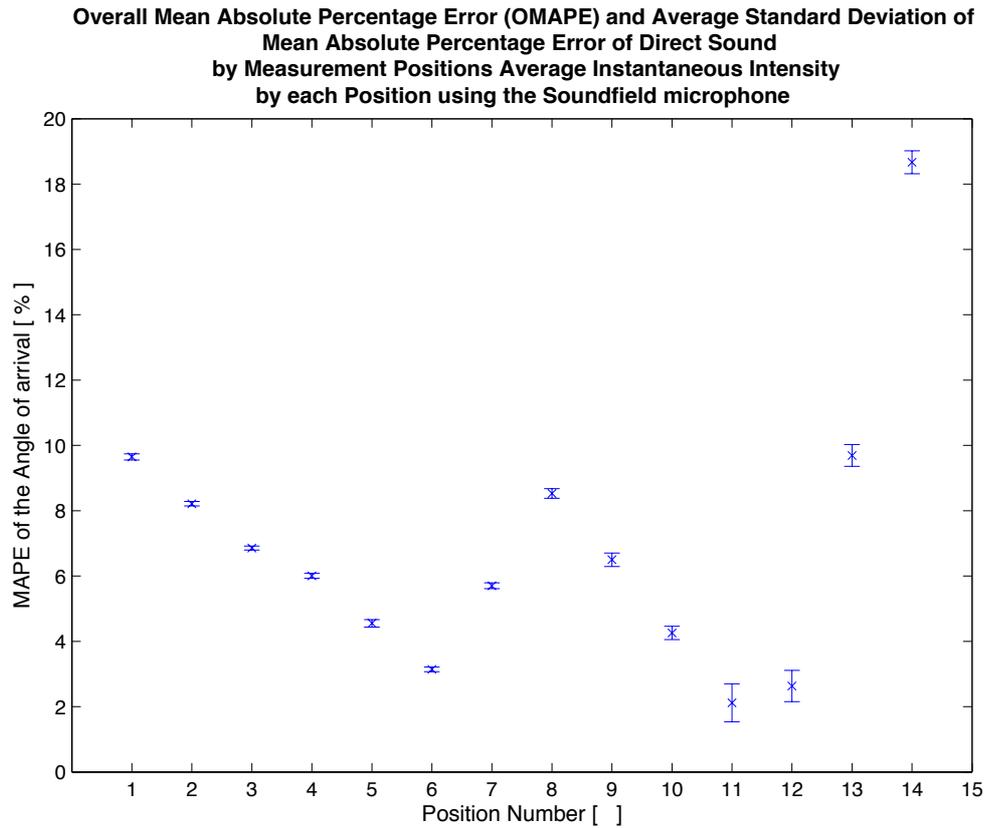
Figure 6.27: Overall mean absolute percentage error ($OMAPE$) of ERB $\left|\overline{\varepsilon}_{\%_{ERB}}\right|$ and overall weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\overline{\varepsilon}_{\%_{ERB}}}$) of direct sound by averaging each instantaneous position (i^{th}) using the ERB's band of the $STFT$ using the *Soundfield* microphone model ST 350.

Overall Mean Absolute Percentage Error (OMAPE) and Average of Standard Deviation of Mean of First Reflection by Measurement Positions Average Instantaneous Intensity by each Position using the Soundfield microphone
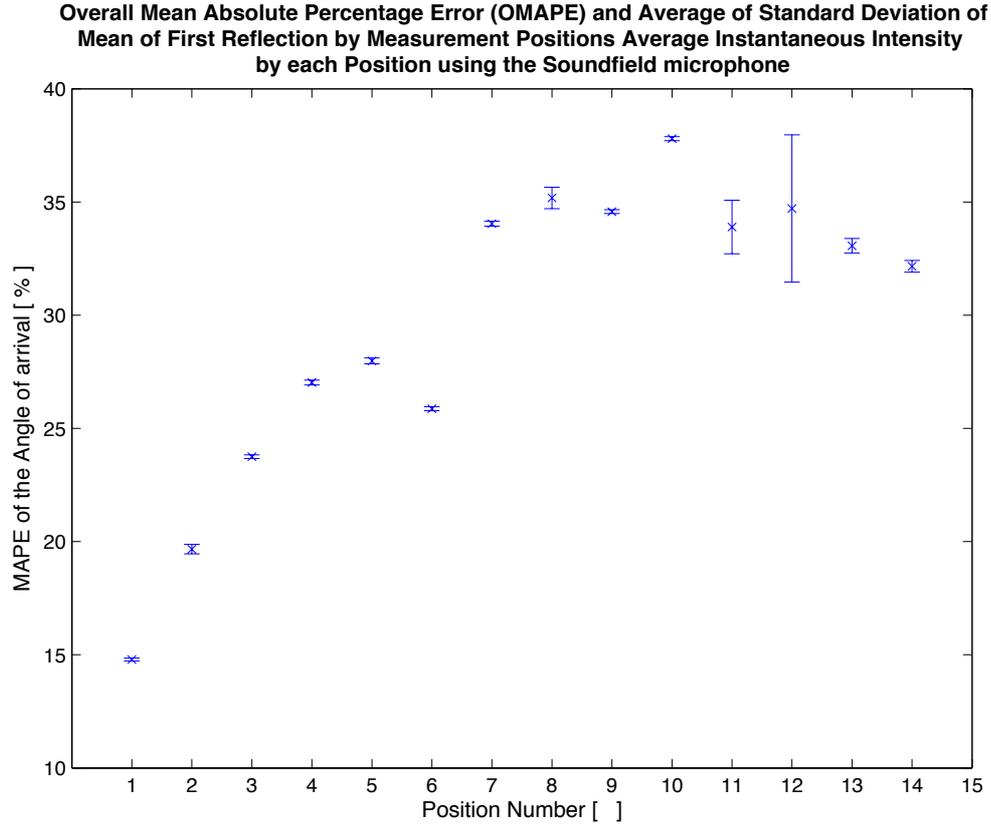
Figure 6.28: Overall mean absolute percentage error ($OMAPE$) of ERB $\left|\overline{\varepsilon}_{\%_{ERB}}\right|$ and overall weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\overline{\varepsilon}_{\%_{ERB}}}$) of first reflection by averaging each instantaneous position (i^{th}) using the ERB's band of the $STFT$ using the $Soundfield$ microphone model ST 350.

1) The direct sound overall mean error ($\overline{\varepsilon}_{\varepsilon}$) for the direct sound is 1.355º, which in terms of overall mean absolute percentage error ($OMAPE$) of ERB $\left|\overline{\varepsilon}_{\%_{ERB}}\right|$ is 6.893% and a overall weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\varepsilon}$) of ± 0.029º, which in terms of weighted average of standard deviation of mean absolute percentage error of ERB ($\overline{\sigma}_{\overline{\varepsilon}_{\%_{ERB}}}$) is ± 0.128%.

2) The overall mean error ($\overline{\varepsilon}_{\varepsilon}$) for the first reflection is 18.791º, which in terms of overall mean absolute percentage error ($OMAPE$) of ERB $\left|\overline{\varepsilon}_{\%_{ERB}}\right|$ is 29.608% and a weighted

average of standard deviation error ($\bar{\sigma}_{\bar{\varepsilon}}$) of ±0.246º, which in terms of overall weighted average of standard deviation of mean absolute percentage error of *ERB* ($\bar{\sigma}_{\bar{\varepsilon}_{\%_{ERB}}}$) is ±0.464%.

The results show the same outcomes on the overall mean values because the same data is used for calculating the average, even though the direction of calculation was performed vertically (by measurement number) instead of horizontally (by *ERB* band). Nevertheless, the overall weighted average of standard deviation value shows a smaller value than reported in the previous results from section 6.1.1.3.

Soundfield microphone	Direct sound		First reflection	
(ST 350)	Overall Mean Error. $\bar{\varepsilon}_{\bar{\varepsilon}}$	Mean std. dev. of Mean Error $\bar{\sigma}_{\bar{\varepsilon}}$	Overall Mean Error. $\bar{\varepsilon}_{\bar{\varepsilon}}$	Mean std. dev. of Mean Error $\bar{\sigma}_{\bar{\varepsilon}}$
	Mean Absolute Percentage Error (*MAPE*) $\left\|\bar{\varepsilon}_{\%}\right\|$	Weighted Average of Standard Deviation of Mean Absolute Percentage Error $\bar{\sigma}_{\bar{\varepsilon}_{\%}}$	Mean Absolute Percentage Error (*MAPE*) $\left\|\bar{\varepsilon}_{\%}\right\|$	Weighted Average of Standard Deviation of Mean Absolute Percentage Error $\bar{\sigma}_{\bar{\varepsilon}_{\%}}$
Instantaneous Intensity in Time domain. $\vec{i}_{inst}$	$7.071°$ $\left(35.724\%\right)$	$\pm 0.207°$ $\left(\pm 1.180\%\right)$	$19.113°$ $\left(31.185\%\right)$	$\pm 2.416°$ $\left(\pm 3.629\%\right)$
Complex instantaneous intensity. $\vec{i}_{complex}$	$8.208°$ $\left(39.630\%\right)$	$\pm 0.330°$ $\left(\pm 1.580\%\right)$	$15.895°$ $\left(24.764\%\right)$	$\pm 1.248°$ $\left(\pm 1.366\%\right)$
Active Instantaneous Intensity with *STFT.* $\vec{I}_{a}$	$1.351°$ $\left(17.280\%\right)$	$\pm 0.013°$ $\left(\pm 0.060\%\right)$	$18.346°$ $\left(30.286\%\right)$	$\pm 0.198°$ $\left(\pm 1.333\%\right)$
Instantaneous Intensity Average each position. $\left\langle \vec{I}_{a} \right\rangle$	$1.355°$ $\left(6.893\%\right)$	$\pm 0.029°$ $\left(\pm 0.128\%\right)$	$18.791°$ $\left(29.608\%\right)$	$\pm 0.246°$ $\left(\pm 0.464\%\right)$

Table 6.1: Summarised results for angular accuracy of direct and first reflection using different techniques performed with the *Soundfield* microphone model ST 350.

6.1.2 Measurements using the one-dimensional p-p intensity probe rotated on orthogonal axis

6.1.2.1 Time domain method using instantaneous intensity ($\vec{i}_{inst}$) using the p-p probe

The following graphs are shown in order to compare the accuracy of angle estimation of direct sound and first reflection using time domain (TD) method with the complex instantaneous intensity ($\vec{i}_{complex}$).

Overall Mean Absolute Percentage Error (OMAPE) and Average of Standard Deviation of Mean Absolute Percentage Error of Direct Sound by Measurement of Positions in Time Domain using the p–p probe
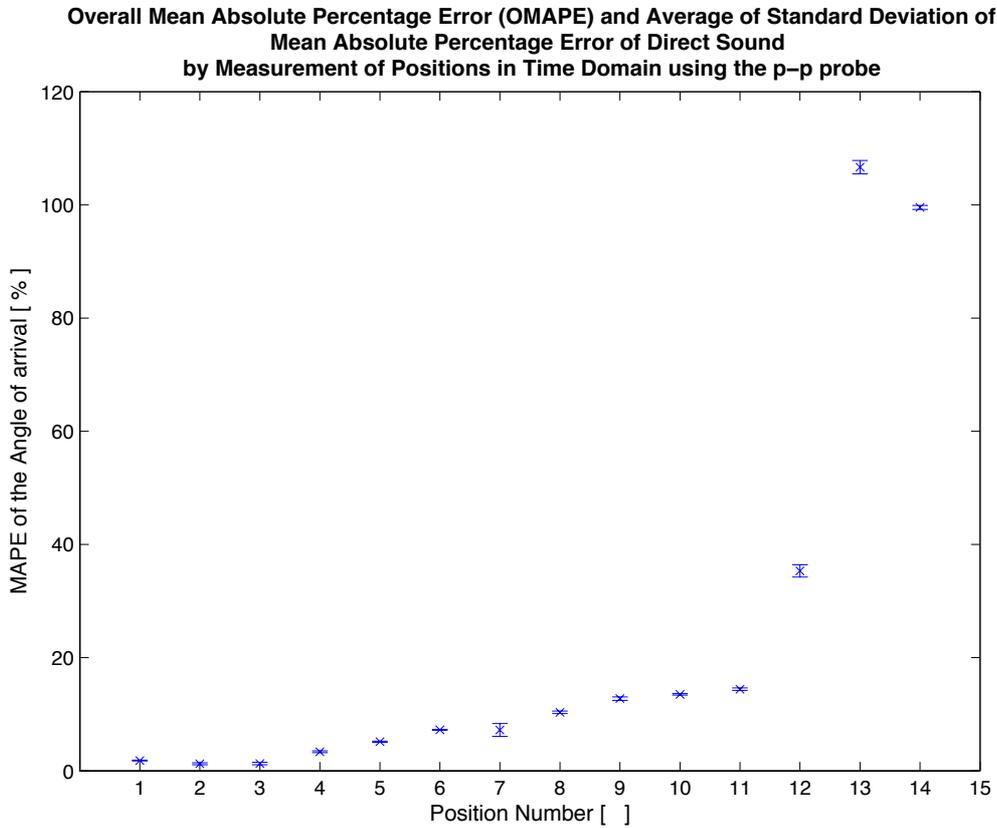
Figure 6.29: Overall mean absolute percentage error ($OMAPE$) $\left|\bar{\varepsilon}_{\bar{\varepsilon}_{\%}}\right|$ and overall weighted average of standard deviation of mean absolute percentage error $\bar{\sigma}_{\bar{\varepsilon}_{\%}}$ of direct sound with instantaneous intensity ($\vec{i}_{inst}$) in time domain (TD) across all (i^{th}) positions using the p-p intensity probe.
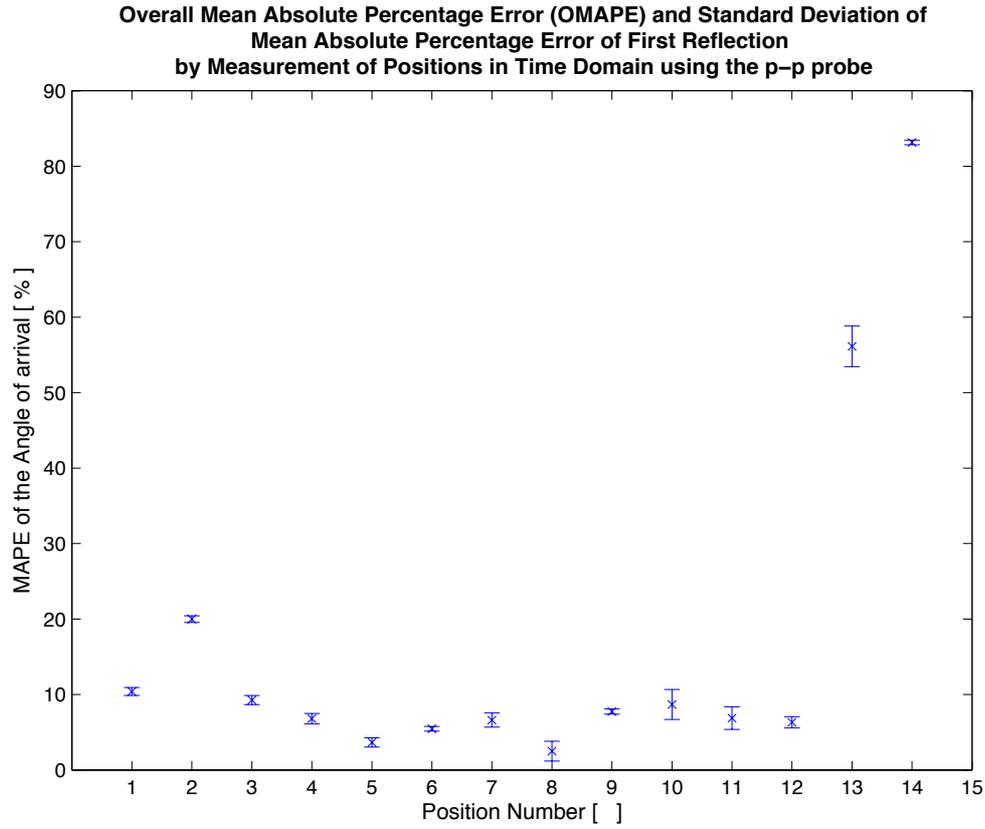
Figure 6.30: Overall mean absolute percentage error ($OMAPE$) $\left|\overline{\varepsilon}_{\overline{\varepsilon}_{\%}}\right|$ and overall weighted average of standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ of first reflection with instantaneous intensity ($\vec{i}_{inst}$) in time domain (TD) across all positions (i^{th}) using the p-p intensity probe.

Using the Instantaneous Intensity gives the following overall results:

1) The direct sound overall mean error ($\overline{\varepsilon}_{\overline{\varepsilon}}$) for the direct sound is 4.467º, which corresponds to an overall mean absolute percentage error ($OMAPE$) $\left|\overline{\varepsilon}_{\overline{\varepsilon}_{\%}}\right|$ of 22.856% and a weighted average of standard deviation error ($\overline{\sigma}_{\overline{\varepsilon}}$) of ±0.066º, which corresponds to an weighted average of standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ of ±0.273% .

2) The overall mean error ($\bar{\varepsilon}_{\bar{\varepsilon}}$) for the first reflection is 3.193º, which corresponds to an overall mean absolute percentage error ($OMAPE$) $\left|\bar{\varepsilon}_{\bar{\varepsilon}_{\%}}\right|$ of 16.683% with a weighted average of standard deviation error ($\bar{\sigma}_{\varepsilon}$) of ±0.319º, which corresponds to an weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) of ±0.565%.

6.1.2.2 Time domain method using the complex instantaneous intensity ($\vec{i}_{complex}$) using the p-p probe

The following graphs show the improvement in accuracy using complex instantaneous intensity ($\vec{i}_{complex}$) to estimate the angle of arrival of the direct sound and first reflection in the case of p-p intensity probe.
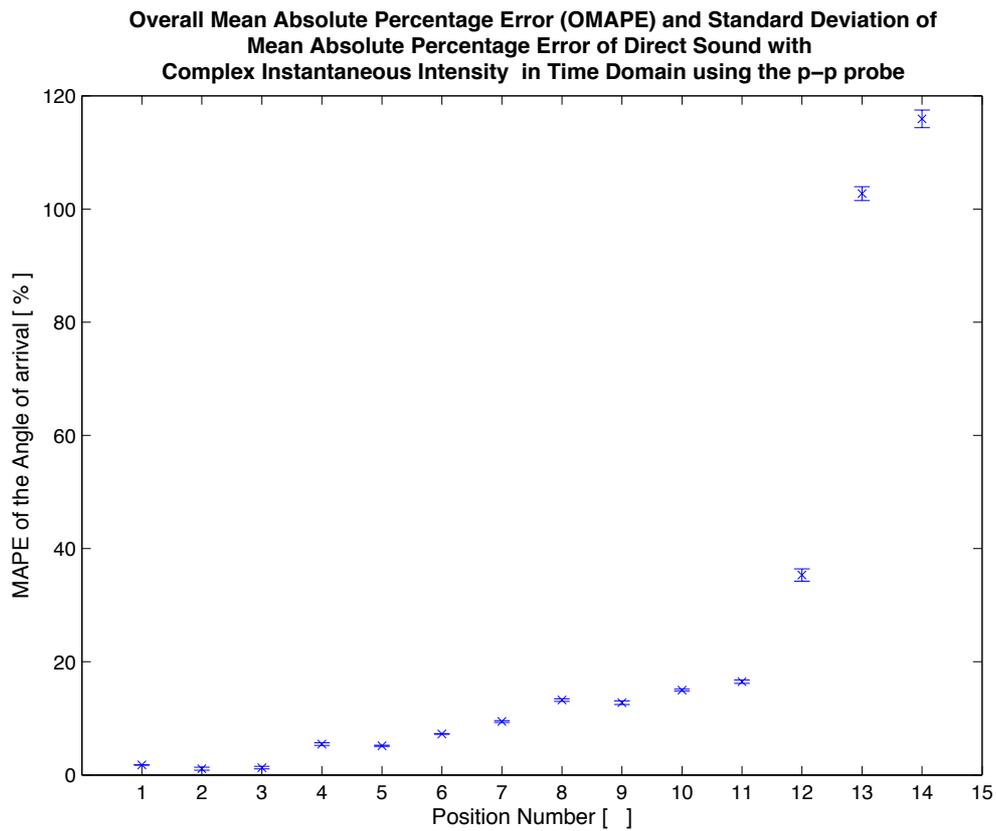
Figure 6.31: Overall mean absolute percentage error ($OMAPE$) $\left|\overline{\varepsilon}_{\overline{\varepsilon}_{\%}}\right|$ and overall weighted average of standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ of direct sound in time domain (TD) with complex instantaneous intensity ($\vec{i}_{complex}$) across all positions using the p-p intensity probe.
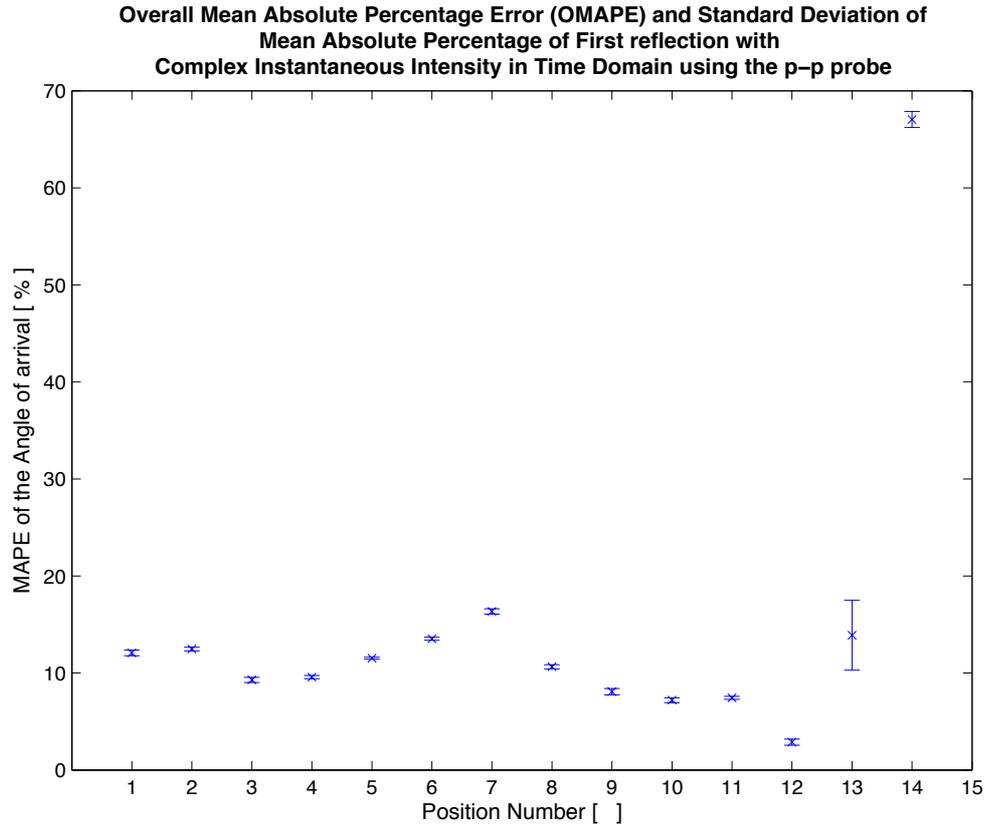
Figure 6.32: Overall mean absolute percentage error ($OMAPE$) $\left|\overline{\varepsilon}_{\overline{\varepsilon}_{\%}}\right|$ and weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) of first reflection in time domain (TD) with complex instantaneous intensity ($\vec{i}_{complex}$) across all positions using the p-p intensity probe.

Using the complex instantaneous intensity ($\vec{i}_{complex}$) gives the following overall results:

1) The direct sound overall mean error ($\overline{\varepsilon}_{\overline{\varepsilon}}$) for the direct sound is 4.838°, which corresponds to an overall mean absolute percentage error ($OMAPE$) $\left|\overline{\varepsilon}_{\overline{\varepsilon}_{\%}}\right|$ of 24.494% and a weighted average of standard deviation error ($\overline{\sigma}_{\overline{\varepsilon}}$) of ±0.057°, which corresponds to an weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) of ±0.312%.

2) The overall mean error ($\bar{\varepsilon}_{\bar{\varepsilon}}$) for the first reflection is 8.582º, which corresponds to an overall mean absolute percentage error ($OMAPE$) $\left|\bar{\varepsilon}_{\bar{\varepsilon}_{\%}}\right|$ of 14.422% with a weighted average of standard deviation error ($\bar{\sigma}_{\varepsilon}$) of ± 0.238º, which corresponds to an weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) of ±0.495%.

6.1.2.3 Short-time Fourier transform ($STFT$) method using active instantaneous intensity ($\vec{I}_a$) using the p-p probe

The following graphs are the mean absolute percentage errors ($MAPE$) and weighted average of standard deviation of mean absolute percentage errors ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) for direct sound and first reflections across the 16 ERB bands using the face-to-face p-p probe using the latest prototype of the custom cradle orthogonal rotation of *x-z* axis.
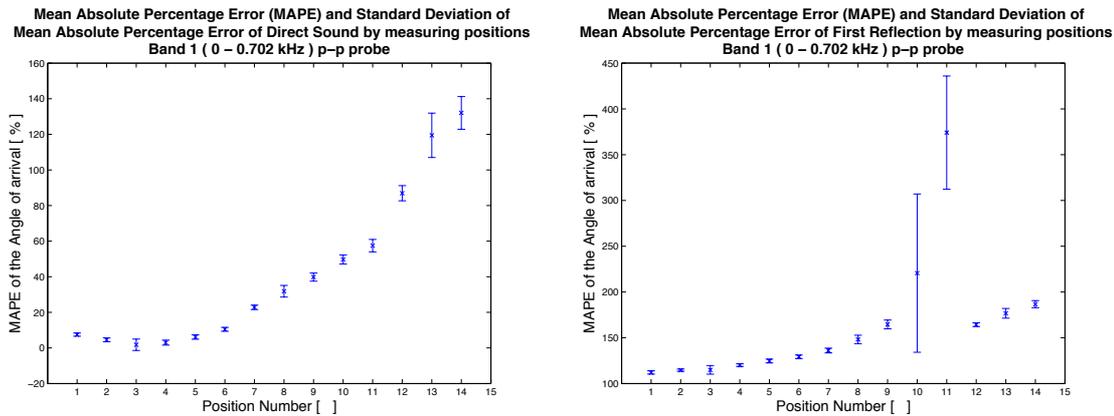
Figure 6.33: ERB band 1 mean absolute percentage error ($MAPE$) $\left|\bar{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\bar{\sigma}_{\bar{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.
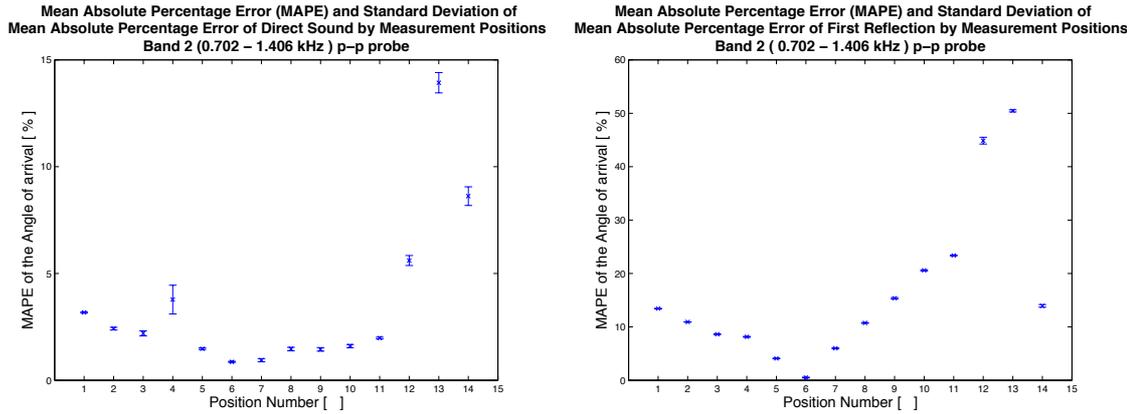
Figure 6.34: *ERB* band 2 mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.
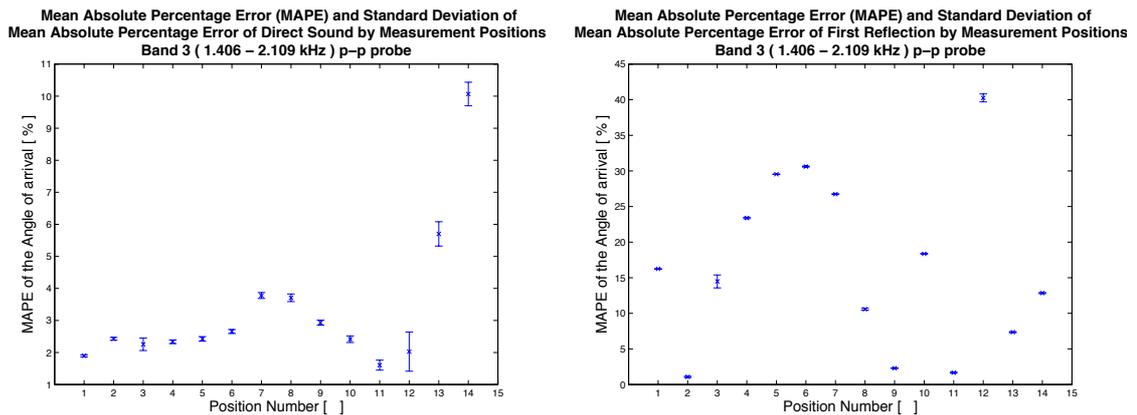
Figure 6.35: *ERB* band 3 mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.
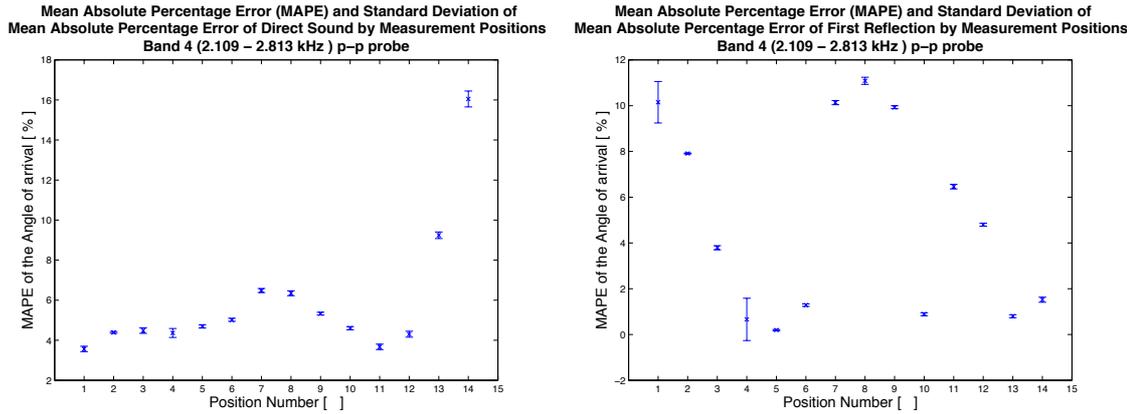
Figure 6.36: *ERB* band 4 mean absolute percentage error (*MAPE*) $\left| \bar{\varepsilon}_{\%} \right|$ and weighted average standard deviation of mean absolute percentage error $\bar{\sigma}_{\bar{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.
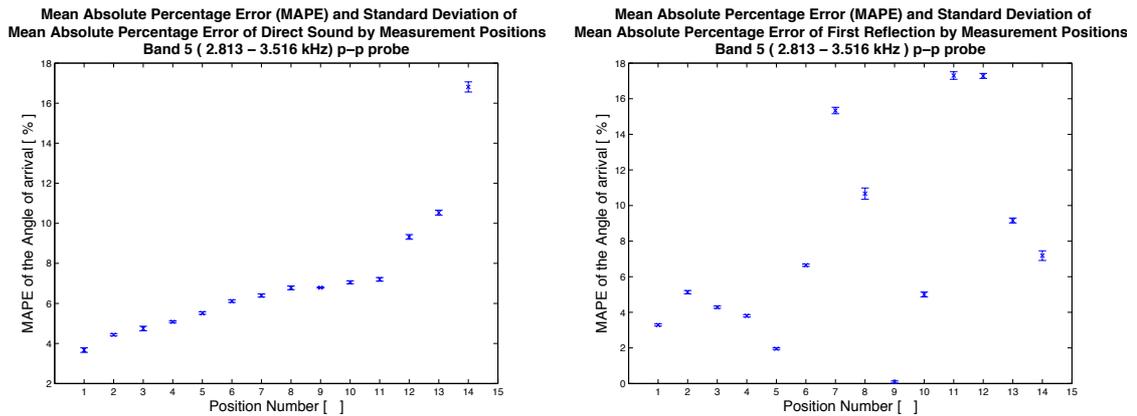
Figure 6.37: *ERB* band 5 mean absolute percentage error (*MAPE*) $\left| \bar{\varepsilon}_{\%} \right|$ and weighted average standard deviation of mean absolute percentage error $\bar{\sigma}_{\bar{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.

141
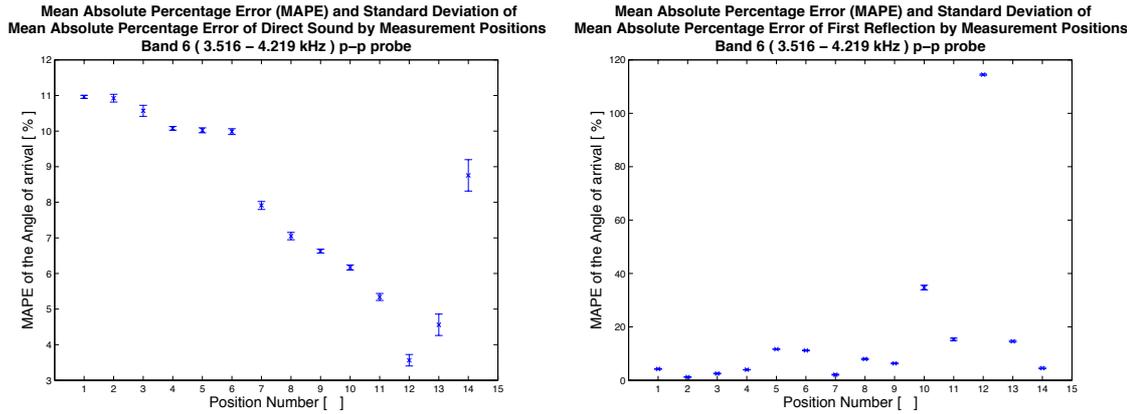
Figure 6.38: *ERB* band 6 mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{rel}\right|$ and weighted average standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.
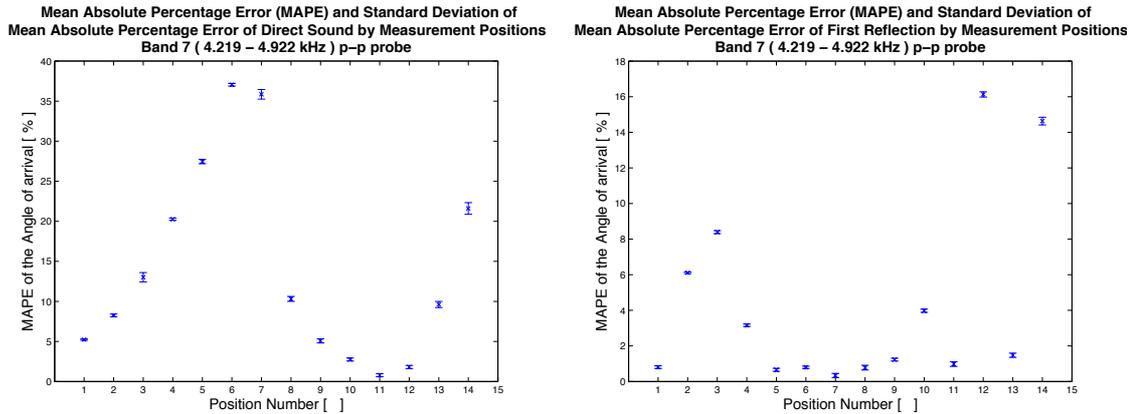
Figure 6.39: *ERB* band 7 mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.

Mean Absolute Percentage Error (MAPE) and Standard Deviation of Mean Absolute Percentage Error of Direct Sound by Measurement Positions Band 8 (4.922 – 5.625 kHz) p–p probe

Mean Absolute Percentage Error (MAPE) and Standard Deviation of Mean Absolute Percentage Error of First Reflection by Measurement Positions Band 8 (4.922 – 5.625 kHz) p–p probe

Figure 6.40: *ERB* band 8 mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.

Mean Absolute Percentage Error (MAPE) and Standard Deviation of Mean Absolute Percentage Error of Direct Sound by Measurement Positions Band 9 (5.625 – 6.328 kHz) p–p probe

Mean Absolute Percentage Error (MAPE) and Standard Deviation of Mean Absolute Percentage Error of First Reflection by Measurement Positions Band 9 (5.625 – 6.328 kHz) p–p probe
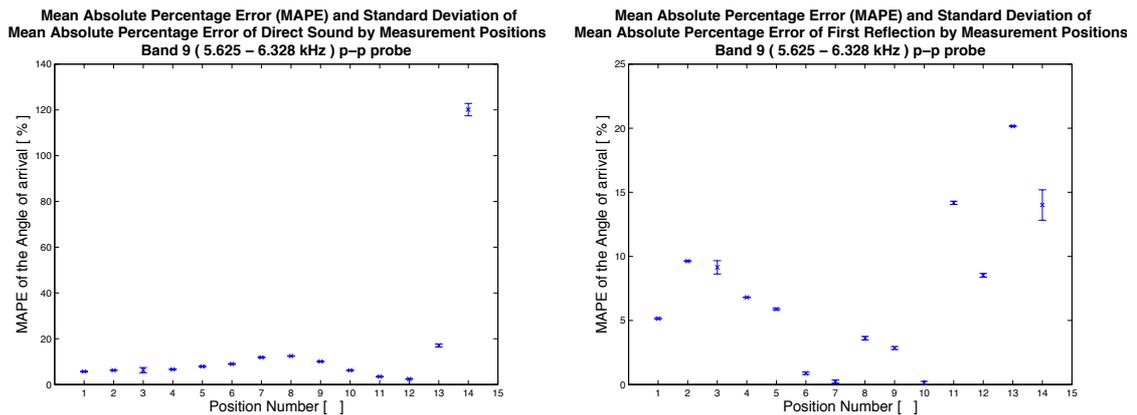
Figure 6.41: *ERB* band 9 mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.
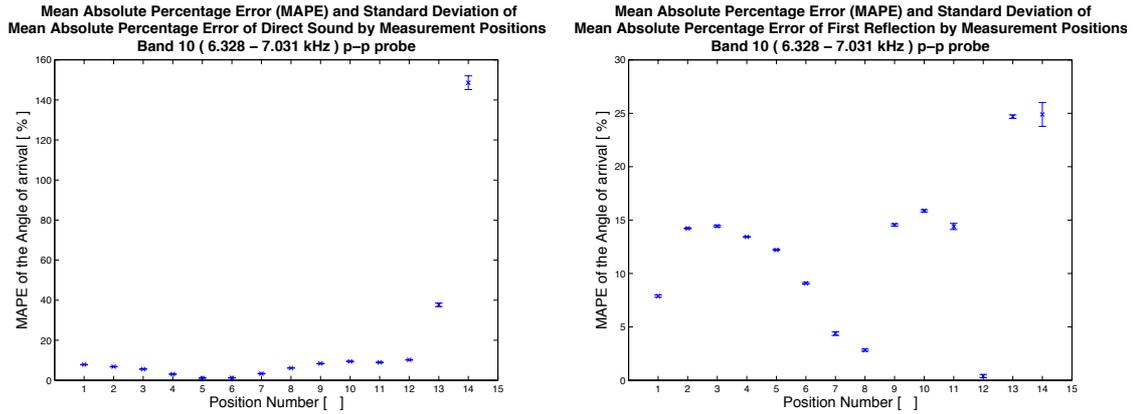
Figure 6.42: *ERB* band 10 mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.
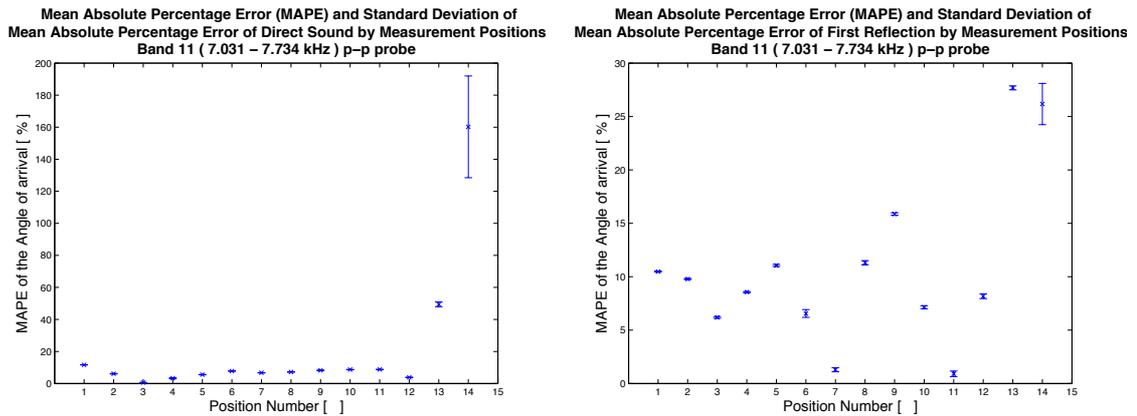
Figure 6.43: *ERB* band 11 mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.
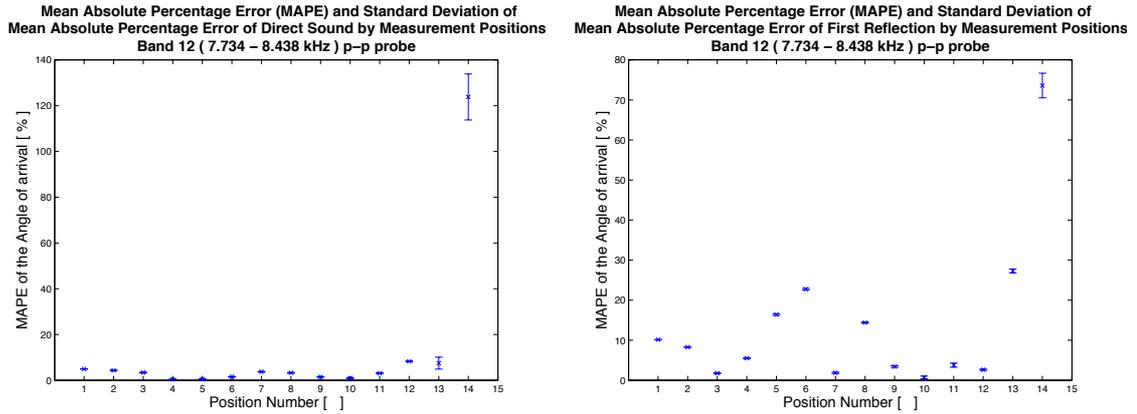
Figure 6.44: *ERB* band 12 mean absolute percentage error (*MAPE*) $\left|\bar{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\bar{\sigma}_{\bar{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.
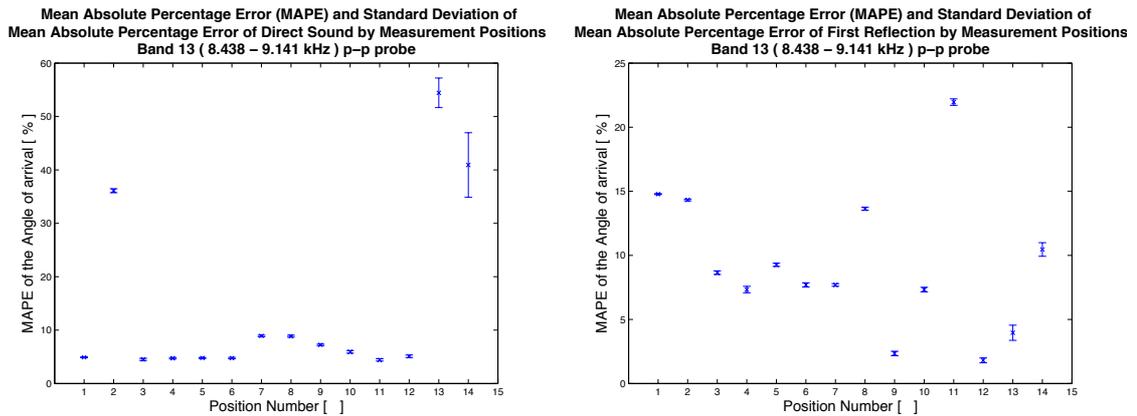
Figure 6.45: *ERB* band 13 mean absolute percentage error (*MAPE*) $\left|\bar{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\bar{\sigma}_{\bar{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.

Mean Absolute Percentage Error (MAPE) and Standard Deviation of Mean Absolute Percentage Error of Direct Sound by Measurement Positions Band 14 (9.141 – 9.844 kHz) p–p probe

Mean Absolute Percentage Error (MAPE) and Standard Deviation of Mean Absolute Percentage Error of First Reflection by Measurement Positions Band 14 (9.141 – 9.844 kHz) p–p probe
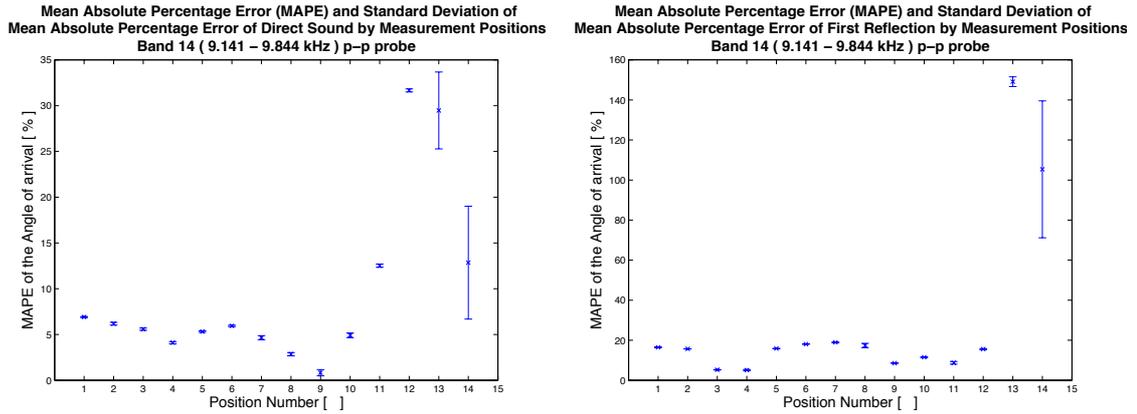
Figure 6.46: *ERB* band 14 mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.

Mean Absolute Percentage Error (MAPE) and Standard Deviation of Mean Absolute Percentage Error of Direct Sound by Measurement Positions Band 15 (9.844 – 10.547 kHz) p–p probe

Mean Absolute Percentage Error (MAPE) and Standard Deviation of Mean Absolute Percentage Error of First Reflection by Measurement Positions Band 15 (9.844 – 10.547 kHz) p–p probe
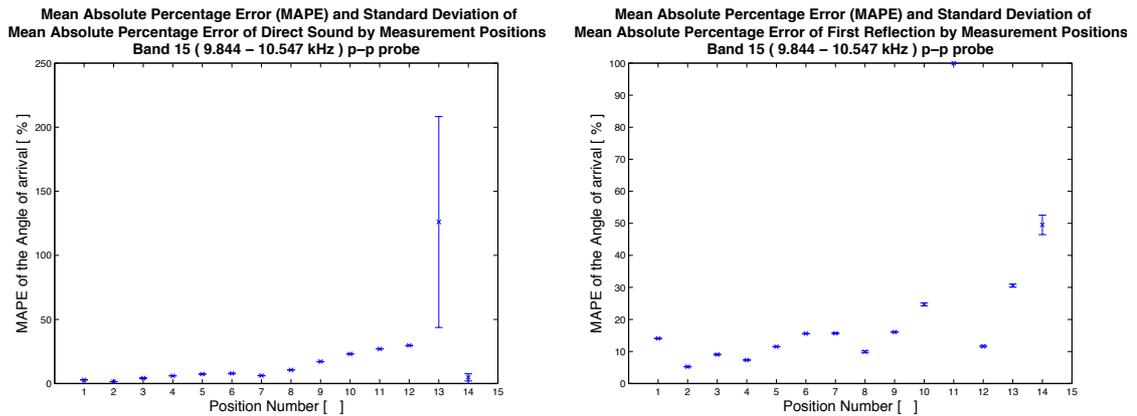
Figure 6.47: *ERB* band 15 mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.
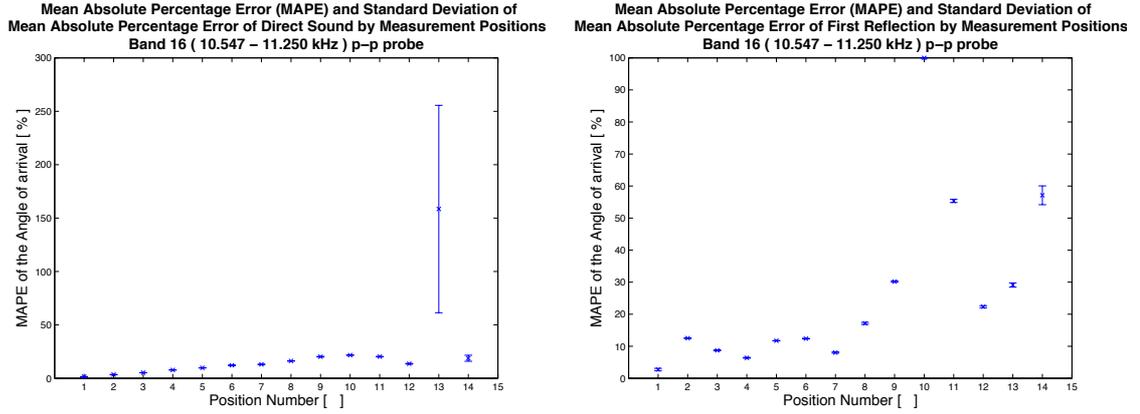
Figure 6.48: *ERB* band 16 mean absolute percentage error (*MAPE*) $\left|\overline{\varepsilon}_{\%}\right|$ and weighted average standard deviation of mean absolute percentage error $\overline{\sigma}_{\overline{\varepsilon}_{\%}}$ for direct sound (left) and first reflection (right) for p-p probe.

The main results are summarised in the following outcomes:

1) The direct sound overall mean error ($\overline{\overline{\varepsilon}}_{\overline{\varepsilon}}$) for the direct sound is 0.163º, which in terms of overall mean absolute percentage error (*OMAPE*) $\left|\overline{\varepsilon}_{\overline{\varepsilon}_{\%}}\right|$ is 14.628% and a weighted average of standard deviation error ($\overline{\sigma}_{\overline{\varepsilon}}$) of ±0.367º, which in terms of weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) represents ±1.092%.

2) The overall mean error ($\overline{\overline{\varepsilon}}_{\overline{\varepsilon}}$) for the first reflection is 4.385º which in terms of overall mean absolute percentage error (*OMAPE*) $\left|\overline{\varepsilon}_{\overline{\varepsilon}_{\%}}\right|$ is 22.909% and with a weighted average of standard deviation error ($\overline{\sigma}_{\overline{\varepsilon}}$) of ±0.788º, which in terms of weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) represents ±0.920%.

After the analysis of the performance of the *Soundfield* microphone by 16 *ERB* bands using the direct sound and the first reflection, it was found that a synthesis of these data in a graph can show the bands that perform better and it is possible to select the frequency working range of the system:
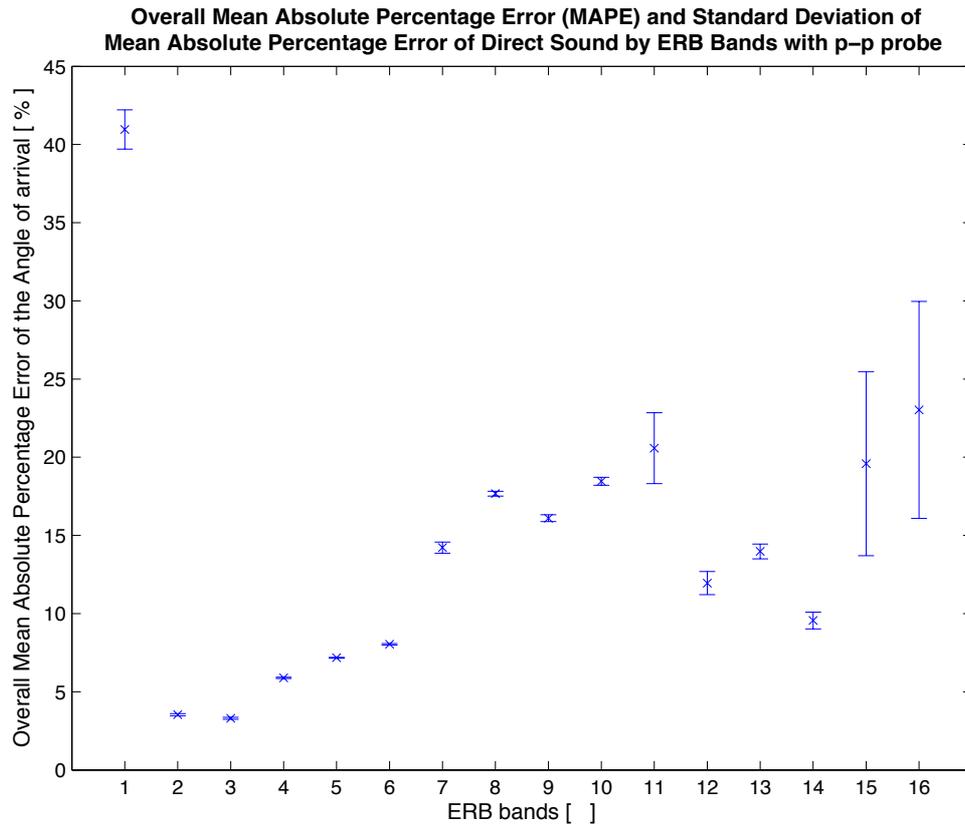
Figure 6.49: Overall mean absolute percentage error (*OMAPE*) $\left|\bar{\varepsilon}_{\%}\right|$ and overall weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) of direct sound by *ERB* bands using the p-p probe.

Overall Mean Absolute Percentage Error (MAPE) and Standard Deviation of Mean Absolute Percentage Error of First Reflection by ERB Band with p–p probe
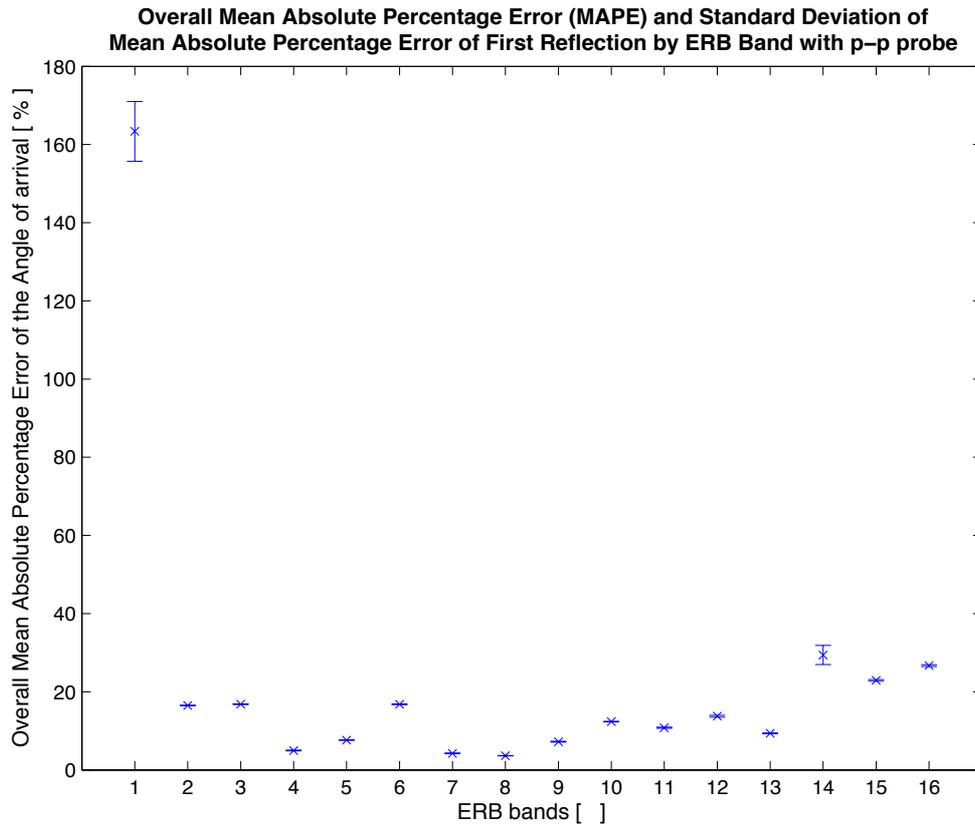
Figure 6.50: Overall mean absolute percentage error ($OMAPE$) $\left|\bar{\varepsilon}_{\%}\right|$ and overall weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) of first reflection by ERB bands using the p-p probe.

6.1.2.4 Average of each position using short-time Fourier transform ($STFT$) using the p-p probe

By virtue of having better estimation of angles using just the average of each measure on all its ERB bands, the same results are presented for the p-p intensity probe.

Overall Mean Absolute Percentage Error (OMAPE) and Standard Deviation of Mean Absolute Percentage Error of Direct Sound by Measurement Positions Average Instantaneous Intensity using the p–p probe
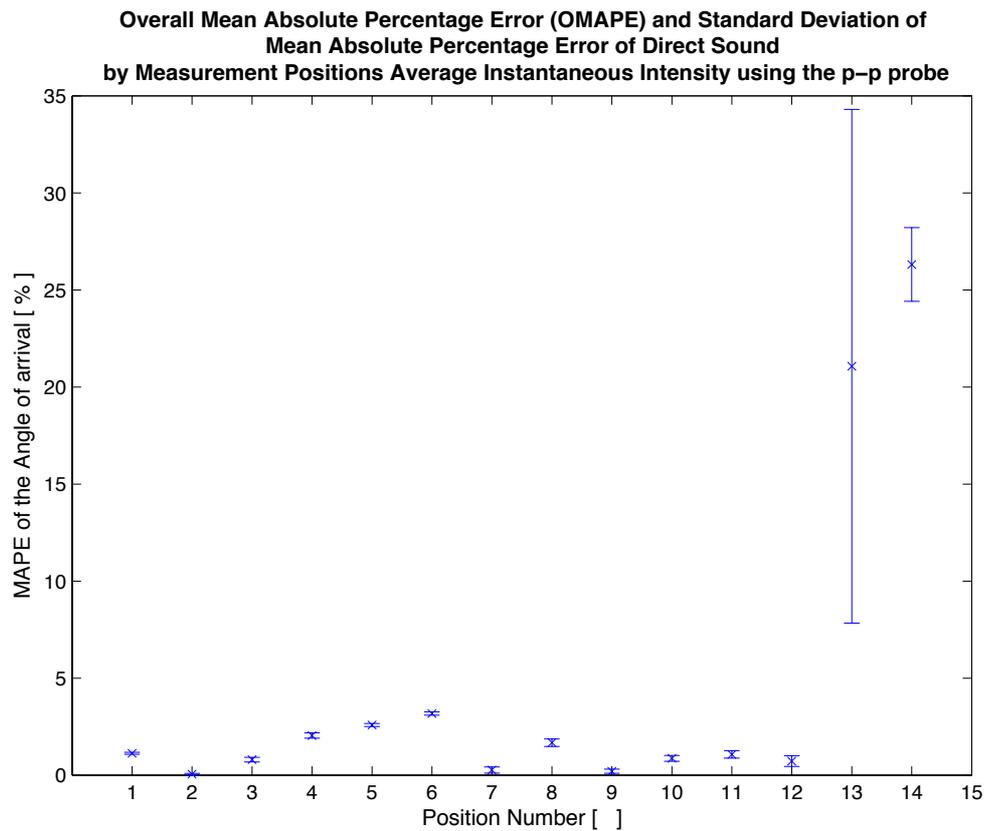
Figure 6.51: Overall mean absolute percentage error ($OMAPE$) $\left|\overline{\varepsilon}_{\overline{\varepsilon}_{\%}}\right|$ and weighted average of standard deviation of mean absolute percentage error ($\overline{\sigma}_{\overline{\varepsilon}_{\%}}$) of direct sound averaging instantaneous intensity ($\vec{i}_{inst}$) by each position using the p-p intensity probe.

Overall Mean Absolute Percentage Error (OMAPE) and Standard Deviation of Mean Absolute Percentage Error of First Reflection by Measurement Positions Average Instantaneous Intensity using the p–p probe
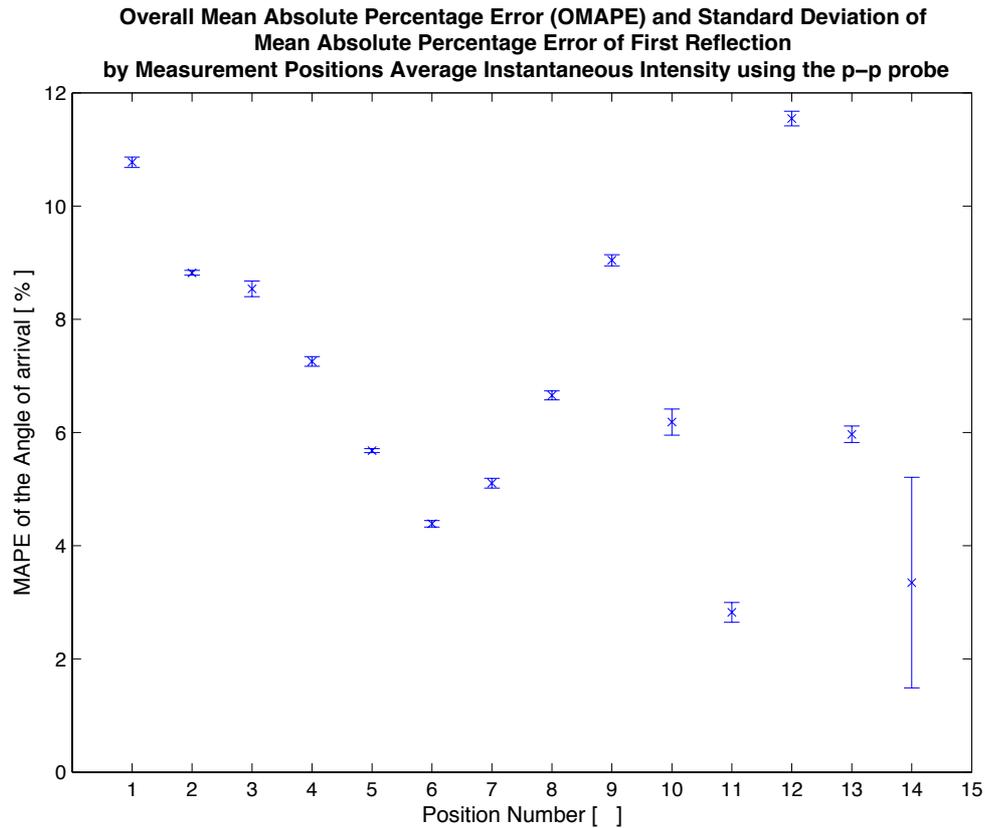
Figure 6.52: Overall mean absolute percentage error ($OMAPE$) $\left| \bar{\varepsilon}_{\bar{\varepsilon}_{\%}} \right|$ and weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) of first reflection averaging instantaneous intensity ($\vec{i}_{inst}$) by measurement (i^{th}) positions using the p-p intensity probe.

1) The direct sound overall mean error ($\bar{\varepsilon}_{\bar{\varepsilon}}$) for the direct sound is 0.226º, which corresponds to an overall mean absolute percentage error ($OMAPE$) $\left| \bar{\varepsilon}_{\bar{\varepsilon}_{\%}} \right|$ of 4.428% and a weighted average of standard deviation error ($\bar{\sigma}_{\bar{\varepsilon}}$) of ±0.293º, which in terms of weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_{\%}}$) of ±1.752 %.

2) The overall mean error for the first reflection ($\bar{\varepsilon}_{\bar{\varepsilon}}$) is 2.971º, which corresponds to an overall mean absolute percentage error ($OMAPE$) $\left| \bar{\varepsilon}_{\bar{\varepsilon}_{\%}} \right|$ of 6.866% with a mean of

standard deviation error ($\bar{\sigma}_{\bar{\varepsilon}}$) of ±0.110º, which in terms of weighted average of standard deviation of mean absolute percentage error ($\bar{\sigma}_{\bar{\varepsilon}_\%}$) of ±0.251%.

6.1.2.5 The best repeatability case using the p-p intensity probe

Based on the smaller overall mean error ($\bar{\varepsilon}_{\bar{\varepsilon}}$) found in the first reflection, the following selection of weighted average of standard deviation is depicted:

1) The weighted average of standard deviation error of direct sound $\bar{\sigma}_{\bar{\varepsilon}}$ = ±0.293º, which corresponds to a weighted average of standard deviation for mean absolute percentage error $\bar{\sigma}_{\bar{\varepsilon}_\%}$ = ±1.752%.

2) The weighted average of standard deviation error of first reflection $\bar{\sigma}_{\bar{\varepsilon}}$ = ±0.110º, which corresponds to a weighted average of standard deviation for mean absolute percentage error $\bar{\sigma}_{\bar{\varepsilon}_\%}$ = ±0.251%.

The algorithm to calculate a weighted average of standard deviation ($\bar{\sigma}_{\bar{\varepsilon}_\%}$), (please refer to equation 3.51 in page 55) does give smaller values in percentage. This is due to the fact that the biggest standard deviation will not dominate the value of the weighted mean because the standard deviations are squared and then summed as variances (σ^2), and then the squared root is taken and divided by the number of samples because calculate the arithmetic mean is not valid for standard deviations.

p-p intensity probe (face-to-face)	Direct sound		First reflection	
	Overall Mean error. $\bar{\varepsilon}_{\bar{\varepsilon}}$	Mean std. dev. of Mean Error $\bar{\sigma}_{\bar{\varepsilon}}$	Overall Mean error. $\bar{\varepsilon}_{\bar{\varepsilon}}$	Mean std. dev. of Mean Error $\bar{\sigma}_{\bar{\varepsilon}}$
	Mean Absolute Percentage Error (*MAPE*) $\left|\bar{\varepsilon}_{\%}\right|$	Weighted Average of Standard Deviation of Mean Absolute Percentage Error $\bar{\sigma}_{\bar{\varepsilon}_{\%}}$	Mean Absolute Percentage Error (*MAPE*) $\left|\bar{\varepsilon}_{\%}\right|$	Weighted Average of Standard Deviation of Mean Absolute Percentage Error $\bar{\sigma}_{\bar{\varepsilon}_{\%}}$
Instantaneous Intensity in Time domain. $\vec{i}_{inst}$	$4.467°$ $\left(22.856\%\right)$	$\pm0.066°$ $\left(\pm1.021\%\right)$	$3.193°$ $\left(16.683\%\right)$	$\pm0.319°$ $\left(\pm2.112\%\right)$
Complex instantaneous intensity. $\vec{i}_{complex}$	$4.838°$ $\left(24.494\%\right)$	$\pm0.057°$ $\left(\pm1.164\%\right)$	$8.582°$ $\left(14.422\%\right)$	$\pm0.238°$ $\left(\pm1.855\%\right)$
Active Instantaneous Intensity with *STFT*. $\vec{I}_{a}$	$0.163°$ $\left(14.628\%\right)$	$\pm0.367°$ $\left(\pm16.332\%\right)$	$4.385°$ $\left(22.909\%\right)$	$\pm0.788°$ $\left(\pm13.777\%\right)$
Instantaneous Intensity Average each position. $\left\langle\vec{I}_{a}\right\rangle$	$0.226°$ $\left(4.428\%\right)$	$\pm0.293°$ $\left(\pm1.752\%\right)$	$2.971°$ $\left(6.866\%\right)$	$\pm0.110°$ $\left(0.251\%\right)$

Table 6.2: Summarised results for angular accuracy of direct and first reflection using different techniques performed with the p-p intensity probe.

From **Table 6.1** and **Table 6.2** it can be seen that the best measure for first reflection is acquired using the p-p intensity probe with the short Fourier transform method and the time averaged instantaneous intensity $\left\langle\vec{I}_{inst}\right\rangle$ used in the spatial impulse response rendering *SIRR* method proposed by Merimaa (Merimaa, 2006). This result is the one that is reported in the abstract of this thesis. The best resolution for the estimation of first reflection is $\bar{\varepsilon}_{\bar{\varepsilon}} = 2.971°$, $\bar{\sigma}_{\bar{\varepsilon}} \pm 0.226°$. Consequently, the complex instantaneous intensity ($\vec{i}_{complex}$) used in time domain is not recommended, and the case of complex instantaneous

intensity ($\vec{I}_{complex}$) applied to the short-time Fourier transform (*STFT*) was not investigated, considering the poor results obtained in time domain.

6.1.2.6 Time Domain (*TD*) and short-time Fourier transform (*STFT*) analysis using the p-p intensity probe rotated to orthogonal axis

The procedure to perform this complementary analysis is to calculate in time domain (*TD*) the intensity quantities such as the instantaneous intensity ($\vec{i}_{inst}$) and the complex instantaneous intensity ($\vec{i}_{complex}$) for each orthogonal *x*, *y* and *z*-axis, then combine them to create the corresponding intensity vector along the desired plane as it was reported in (Romero-Perez, 2010). For the p-p intensity probe, this is found in the *Matlab* m file named:

`Extract_winMLS_signals_Corrected69LOG_BFormat_192kHz.m`

In the case of the *Soundfield* microphone model ST350 the script used to analyse the signals is named:

`Extract_winMLS_signals_Corrected70LOG_BFormat.m`

This m files produce a lot of graphs that are useful for diverse analysis of the 3-D impulse response. In this report, only the four most relevant types of graphs are presented.
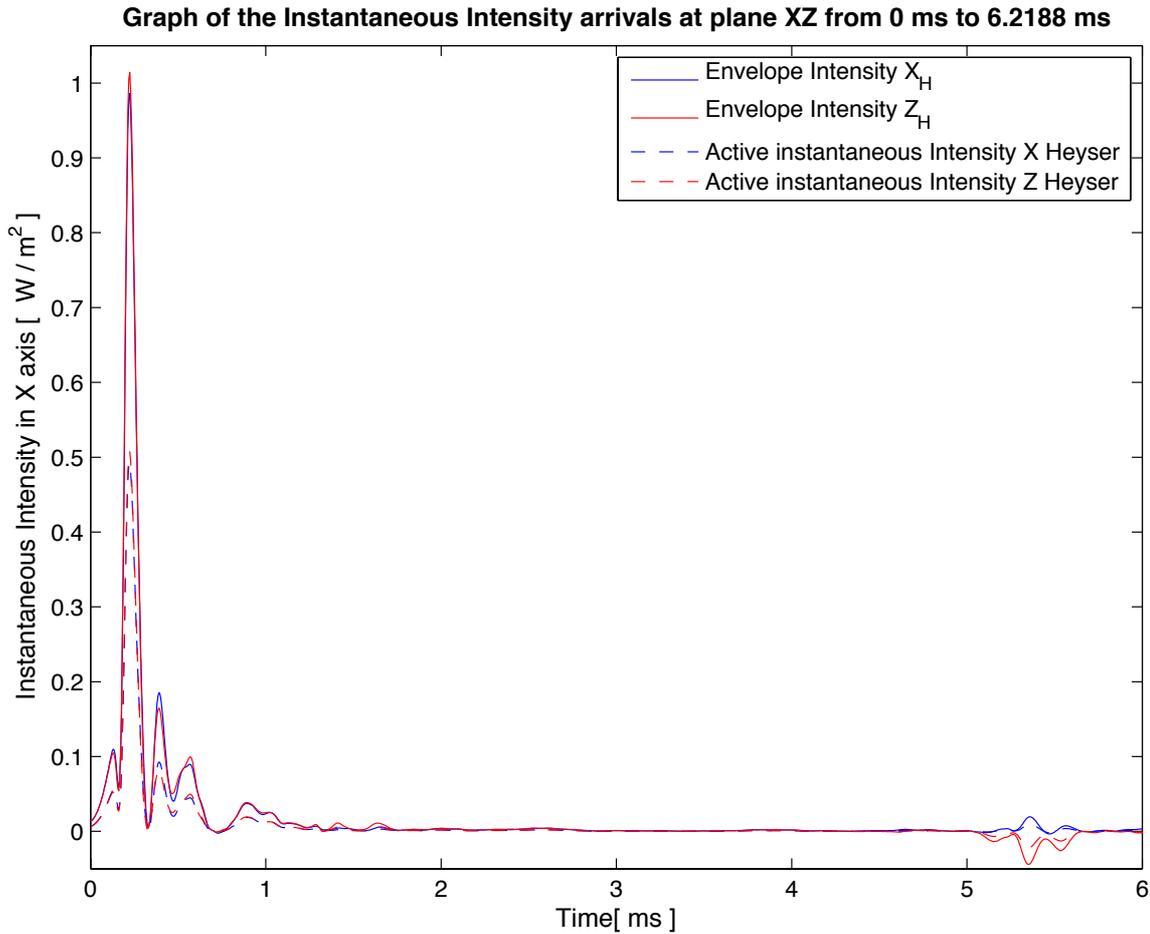
Figure 6.53: A comparison among the components of the envelope of intensity in *x* and *z*-axis ($\overline{i}_x$, $\overline{i}_z$) vs. the active complex instantaneous intensity in *x*-axis ($\vec{i}_{active \atop complex_x}$) and z-axis ($\vec{i}_{active \atop complex_z}$).

In the following example shown in **Figure 6.53** it is possible to see how the different types of intensity are compared for the accuracy detection of the first reflection. It depicts the envelope of intensity ($\overline{i}$) used for detection of direct sound and first reflection when a p-p intensity probe is used in position P_1 measurement 1. It is calculated using equation 3.13 from page 38. It can be seen that the active complex instantaneous intensity components ($\vec{i}_{active \atop complex_x}, \vec{i}_{active \atop complex_z}$) have positive and negative excursions and the envelope intensity components ($\overline{i}_x, \overline{i}_z$) preserve the polarity of them. The active complex instantaneous intensity is calculated using equation 3.33 on page 46. The envelope of a

signal is a positive quantity and may be presented in a logarithmic amplitude scale in order to depict a huge range in the time domain (TD). The envelope is related to the Energy-Time Curve, or *ETC*, known from Time Delay Spectrometry (Thrane et al., 1984). However, it shows more contrast than the instantaneous intensity ($\vec{i}_{inst}$) preserving important details of the signal, which are useful for location of reflections.

In **Figure 6.54** it can be seen that the instantaneous intensity ($\vec{i}_{inst}$) tends to show the instantaneous incoming energy in a way that does not take into account any averaging. Therefore, the localisation of reflections is more difficult, due to the fact that the arrows are more spread than in the case of the use of complex instantaneous intensity ($\vec{i}_{complex}$) as it can be seen in **Figure 6.55**.
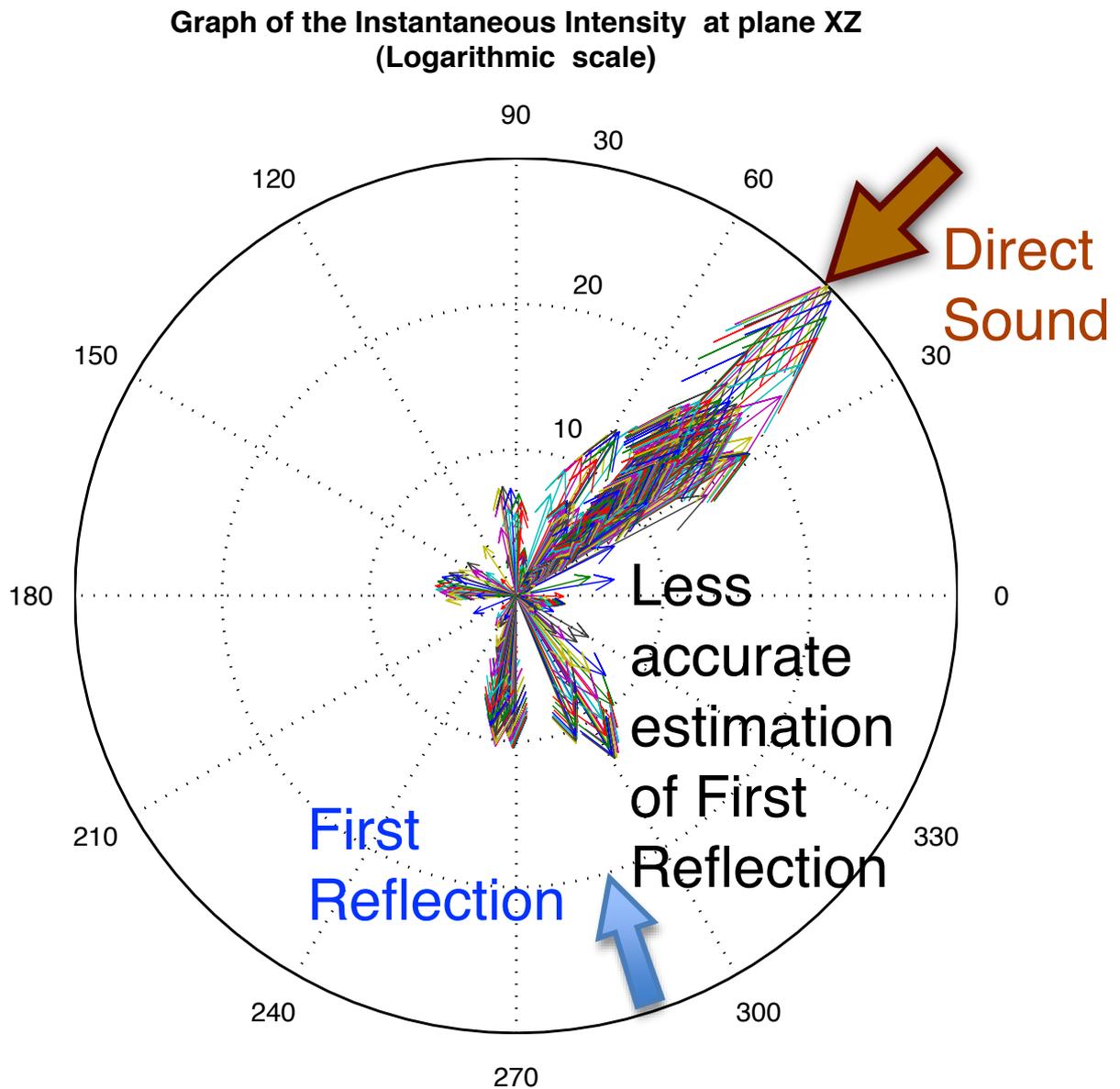
Figure 6.54: Graph of the instantaneous intensity ($\vec{i}_{inst}$) of the whole sound decay using p-p intensity probe, where the blue solid arrow is the real direction of the reflection. Analysis performed under the time domain method.

A comparison of **Figure 6.54** and **Figure 6.55** shows that calculations of active instantaneous intensity ($\vec{i}_a$) with the particle velocity estimate ($\vec{u}$) have a slight deviation on the estimation of the direct sound and its first reflections. After doing the analysis throughout the sample of ten measurements per position, it was found there were some differences in terms of distribution of reflections in time. In terms of the direction of the

direct sound, both techniques converge to the same value. In the case of the first reflection, it is not the same, likely because of the mixing time process that tends to create some uncorrelated sources of patterns of reflections. Another cause of discrepancies on different time of arrival is attributed to the two drivers on the monitor speaker. Both drivers are located in different coordinates; therefore, the location of the direct sound and first reflection may vary slightly in time and direction of arrival (DOA). The other concern is with the different polar directivities that exhibit each driver across frequency. At high frequencies, the tweeter tends to be more directional than at mid frequencies.

As can be seen in **Figure 6.55**, the effect of using the active complex instantaneous intensity in time domain ($\vec{i}_{active \atop complex}$) is an evident aid for detection of the first reflection. It helps to locate it in a more clear way than using the active instantaneous intensity in the time domain ($\vec{i}_{a}$).

**Graph of the Active Complex Instantaneous Intensity at plane XZ
(Logarithmic scale)**
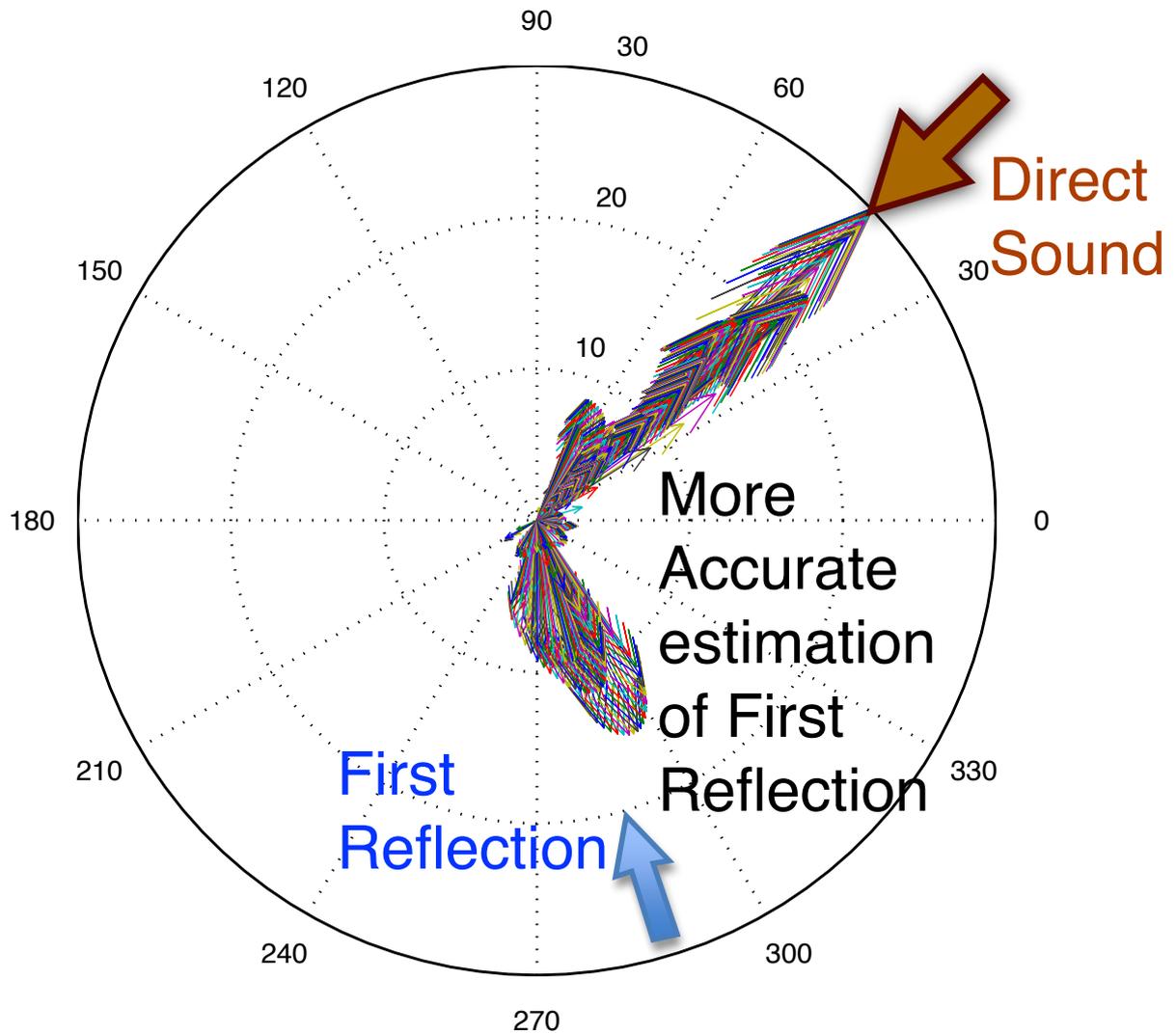
Figure 6.55: Graph of active complex instantaneous intensity ($\vec{i}_{active\ complex}$) of the whole sound

decay using p-p intensity probe, where the blue solid arrow is the real direction of the reflection. Analysis performed under the time domain method (TD).

In the case of the short-time Fourier transform ($STFT$) method and the mean active intensity in frequency domain averaged inside each time window every five samples.
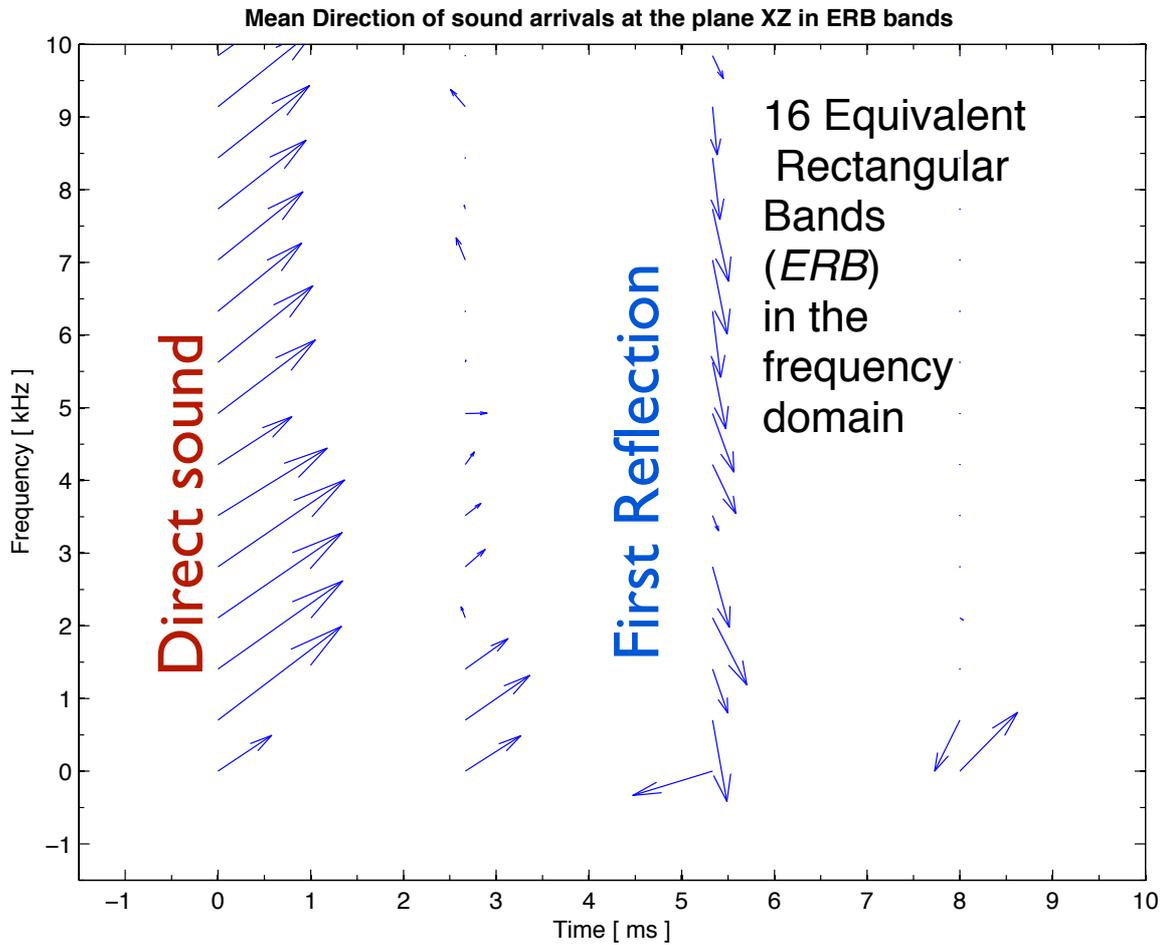
Figure 6.56: Example of short-time Fourier transform ($STFT$) method applied to the p-p intensity probe with vibration damping in position P_1 also known as the calibration position for direct sound ($\theta_{s_1} = 45°$) and first reflection ($\theta_{m_1} = -75.067°$).

Referring to **Figure 6.56** it can be seen that it shows the strength and the direction of the direct sound and the first reflection for position P_1 measurement 1. The length of the Fourier transform ($NFFT = 2048$) and the sample rate is 192 kHz. The number of ERB is 16 and covers up to 10 kHz. The length of the time window is 2.667 ms, overlapping at 50% of its length. The direct sound is located at $t = 0$ and the first reflection is found around 5.385 ms.

The diffuseness estimate (ψ) is a very useful measure for detection of transients in a sound decay. It can be used for detection of specular reflections.
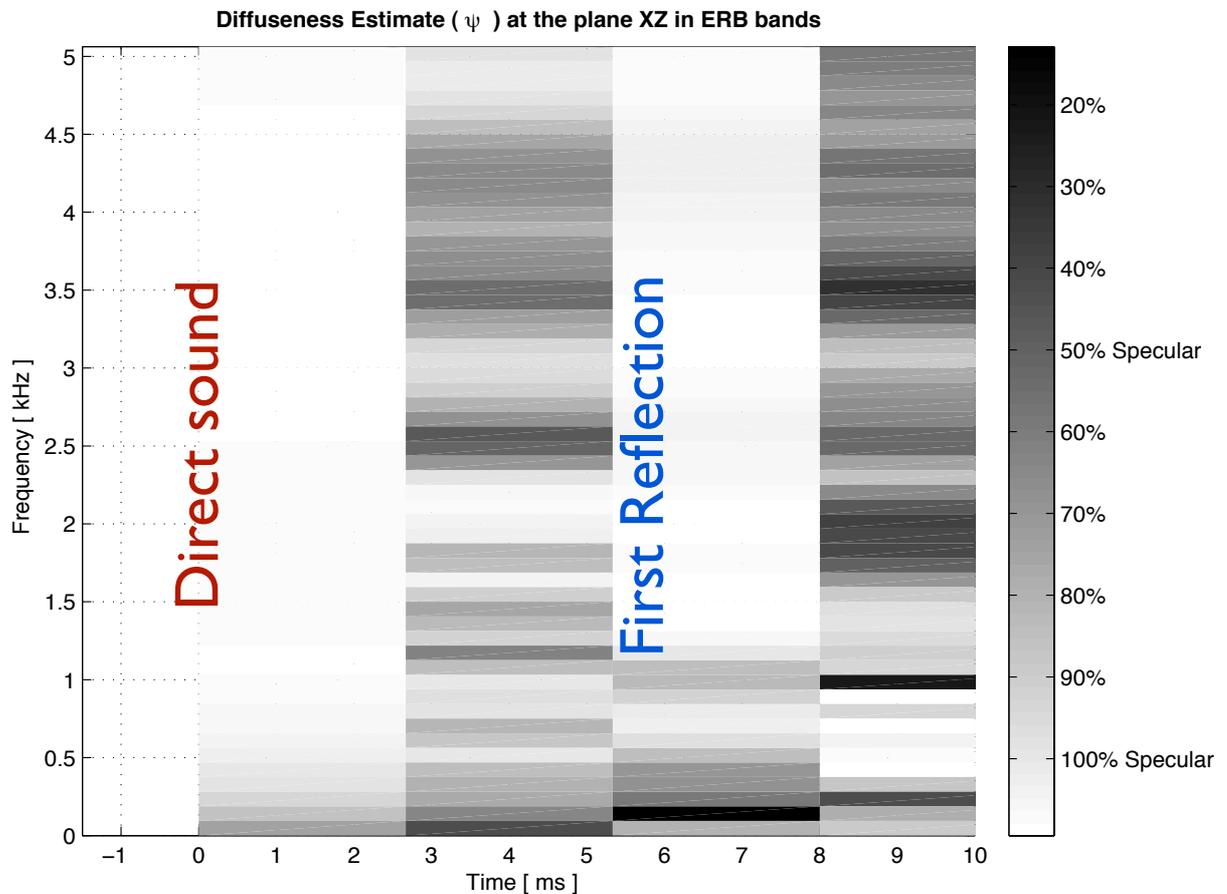
Figure 6.57: Example of diffuseness estimate (ψ) graph made in position 1 measurement 1 with the p-p intensity probe.

6.2 Summary

This section describes the results and discussion. The calculations implemented are described and an adequate number of graphs show the characterisation of the reflections and the measurement of uncertainty on the measurements of direct sound and first reflections. The first topic is the description of the measurement environment, the semi-anechoic chamber, and the monitor speaker influence. Subsequently, the measurement positions and the calculation of the angles with the geometrical are presented. The accuracy of the methods of analysis for all cases is discussed. A table of results is presented in order to compare effects of different probes. For the single reflection scenario, the graphs of the *Soundfield* microphone are depicted. An example of the peak

detection on the B-Format signals is shown. The next graphs are the p-p intensity probe rotated to orthogonal axis. An example of the time domain method (TD) using the envelope of the signals is presented. Successively, a graph of instantaneous intensity ($\vec{i}_{inst}$) is compared with complex instantaneous intensity ($\vec{i}_{complex}$) and the best estimation is found with the first ($\vec{i}_{inst}$). The following graph described is the short Fourier transform method ($STFT$) applied to the p-p intensity probe. The last result presented in this chapter is depicted with the diffuseness estimate (ψ) graph using the p-p intensity probe.

Chapter 7 : Results of practical applications

This chapter aims to investigate the extraction and analysis of temporal and spatial distribution early sound decay in a real room using the *Soundfield* microphone (please refer to section 5.5). A method based on B-format signals is adapted to small rooms. It can map the spatial and temporal distribution of sound energy and diffuseness in three dimensions. The main acoustic problem found in small rooms is coloration of sound. It is perceived as a change of timbre, rhythm sensation and signal pitch and is the main problem encountered when music is reproduced in such spaces (Rubak and Johansen, 2003). It happens in a complicated way, as it depends on temporal and spatial variables. In order to minimize these problems, acoustic treatment needs to be applied to the room. The traditional ways of measuring rooms to assess this treatment had not taken into account the directionality of the sound energy and their effect in terms of perception of stereo image without blur estimation caused by harmful reflections, nor had take into account the characteristics of neutral rooms. This technique is practical to be used to assess 3-D impulse responses (3-D *IR*) of rooms.

7.1 Three-dimensional impulse response (3-D *IR*) measurements in small rooms

These types of practical measurements were performed using a *Soundfield* microphone and a *Focusrite Saffire* 26 i/o soundcard connected to a portable computer. The measurement software performs the extraction of the impulse response of the four B-Format signals. (Refer to **Figure 4.12**). *Matlab* is used for further analysis as discussed in section 4.4. The excitation source was the existing studio monitors found in the rooms.

The impulse response of the room (*IR*) represents the interaction of the speaker and the room at this particular position. It is important to note that these impulse responses are

not good to calculate any acoustic parameter due to the fact that the studio monitors are not omni-directional at all frequencies. To obtain the impulse response, a deconvolution process is performed with the aid of the measurement software *WinMLS*[17]. The impulse response from the measurement is generated in binary or `ACII` file formats. The excitation signal chosen in the software was linear swept sine. To improve accuracy, the sweep length was selected according to the measured *RT*. The four impulse responses were obtained by performing simultaneous *RT* measurements by selecting the '4 channel mode', with the soundcard operating at its highest resolution, which was a bit depth of 24 bits and a sampling rate of 96 kHz. The final step is to export data to *Matlab* for further analysis using `Loadimp.m,` a script provided by *WinMLS*. The B-Format signals are normalized by the maximum value found in the four signals.

Some measurements of a 3-D *IR* in small rooms were done using the *Soundfield* model SPS 422B microphone[18] and applying the short Fourier Transform (*STFT*) method. **Figure 7.1** and **Figure 7.2** show some results of this measurement scenario.

Only the horizontal plane (*xy*-axis) is investigated at this time. Measurements were performed at two different source positions. The first was at the front of the microphone and the second at 120º from the origin, referenced at front with an approximated angle of 0º (which later was measured and was 2º, see **Figure 7.1** and **Figure 7.2**).

The time scale begins at around 0 ms, the useful region is adapted to observe the direct sound and its decay. Looking at **Figure 7.2** in the upper graph, it can be seen that the variation of direction of the early reflections is not traceable after the time labelled around 10 ms. It corresponds to the 10-15 ms limit where the reflections leave the discrete zone and increase its density across time. It is interesting to see that the more diffuse the reflection, the less coherent the direction estimate is across the audible frequency range. In **Figure 7.2** it can be seen that the diffuseness values (ψ), which appear as grey spots, tend to randomly vary after the appearance of the early reflections. Therefore, this is not a useful descriptor of the stage of decay, unless it is combined with the amplitude level of the reflections calculated with the *Matlab*'s `quiver` graph function, which shows the

---

[17] http://www.winmls.com [Online accessed in Oct 20th 2009]
[18] http://www.soundfield.com/products/sps422b.php [Online accessed in Oct 20th 2009]

mean direction of the reflections calculated after smoothing with the *ERB* bank filter. From this analysis, it is clear that new measures for taking into account the mixing time (t_{mix}) and the beginning of the reverberant tail are needed in order to perform an automatic analysis of each room measured. The use of higher-order statistics and modern time series theory may possibly be adapted to cover this needs.

7.1.1 3-D Graphs of short-time Fourier method (*STFT*) using the SPS 422B *Soundfield* microphone
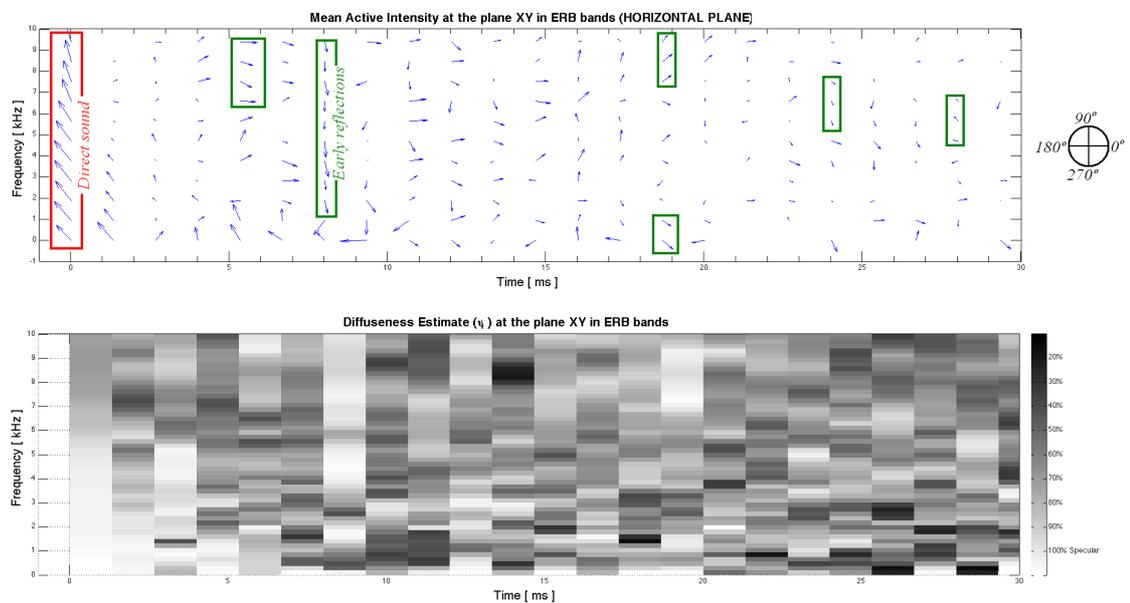
Figure 7.1: Upper graph shows the direct sound, early reflections, and late sound decay in Blue Room 2, which is a small control room at the University of Huddersfield. Lower graph shows the diffuseness estimate (ψ) calculated with the *STFT* method. The source was located at 121º. Note that the front direction is located at 0º at the right side.
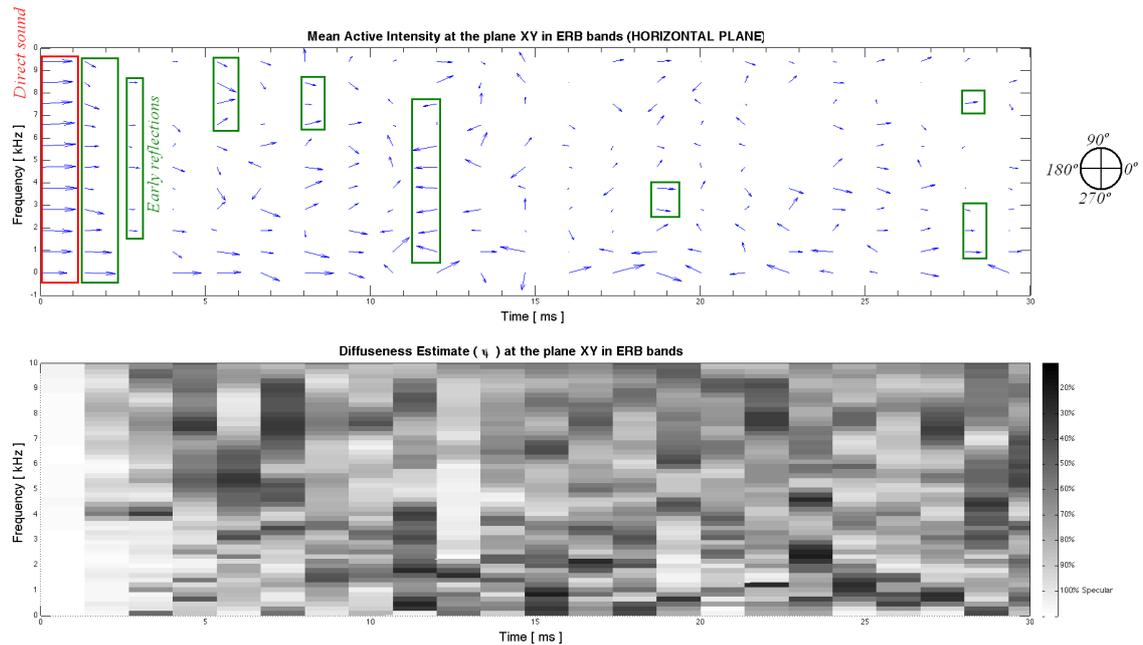
Figure 7.2: Graphs showing the direct and early reflections and late sound decay (upper graph) and the diffuseness estimate (ψ) in the (lower graph). The monitor direction used is from the position 2 (at the front of the microphone). Note that the front direction located at 0º is rotated -90º in the sketch of the control room.

For more details on the implementation of this measurement system, please refer to Romero-Perez (Fazenda and Romero-Perez, 2008, Romero-Perez et al., 2009).

Graph of peak detection applied to B-Format signals
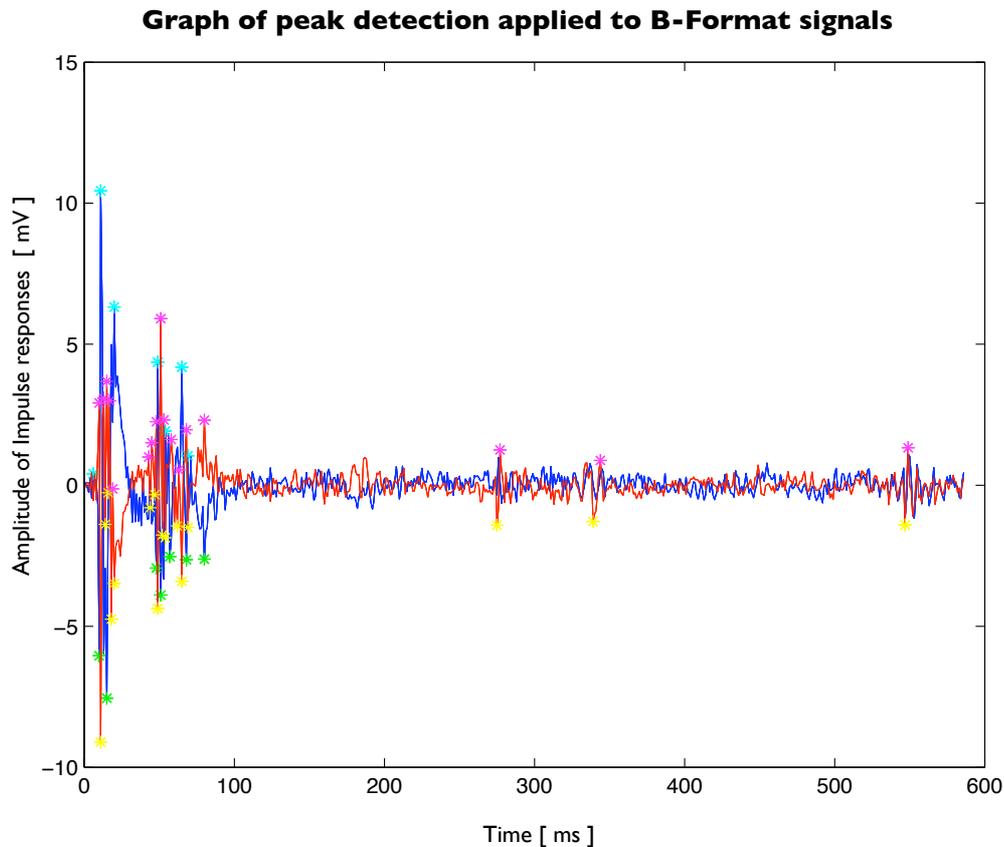
Figure 7.3: Example of peak detection algorithm applied to the B-format signals.

The example of a peak detector application is depicted in **Figure 7.3**. It shows the capabilities of the peak detector algorithm tested. It uses the raw data of B-Format signal to display the peaks and valleys in the signal as asterisks. It can find the peaks either in positive or negative sides with different asterisk colours, and if combined with the right procedures such as using it on the envelope of the intensity, it already can automatically map the early reflections. This is because the procedure of calculating the instantaneous intensity acts like a good filter of relevant information on the signal, and if a much cleaner signal is needed to locate the relevant reflections, then its envelope can simplify the signal furthermore.

7.2 Summary

The estimation of direction of arrival of early reflections using the *STFT* approach in time-frequency domain appears to be a good technique with high accuracy and repeatability in comparison with the time domain analysis. The inclusion of Equivalent Rectangular Band bank filter and localization blur perception limits in the model helps bridge the gap between the objective measures and the subjective measures. This paper reports advances of the research up to date and possible research directions.

The physical limitation of the system is the resolution of the *Soundfield* microphone, especially in the forward and backward directions where the human localisation blur angle is more accurate. The future use of the *Microflown* probe may improve the measurement results, and better prepare the impulse response to create accurate auralisation.

Chapter 8 : Discussion of results

Angles of arrival of sound for their nature exhibit particular problems in error analysis, since in many situations, the "zero degree direction" is arbitrarily chosen and, therefore a percentage error ($\varepsilon_{\%}$) is meaningless. Likewise, an "absolute" angle uncertainty is meaningful, but the percentage of uncertainties feel a rapport, merely when they are related to a difference between two different angles of arrival, or when there is an undoubtedly implicit "zero reference." For example, relating to a full sequence of 360°, where there may be a rotation of a particular reflection in the time series of a sound decay. This is not relevant in the case of an acoustic measurement unless a diffuseness estimation measurement (ψ) is made to treat a room, which has dissimilar absorption coefficients (α_i) and the sound energy is creating some rotation of the reflection patterns across the decay time.

The procedure to obtain the weighted average of the standard deviation as described in the last part of section 3.2.3.3 was tested and it was found that the biggest standard deviation of a particular sample (σ) had more weight in the average because when it was squared to obtain the variance (σ^2) it grew exponentially, not linearly. Therefore, the weighted average of standard deviation ($\bar{\sigma}_{\bar{\varepsilon}}$) was chosen over the mean of the standard deviation. The correction of the small sample using the Student's t-distribution increases a factor of 1.833 of the original calculated value of the weighted average of standard deviation. The best repeatability case was found to be with the p-p probe, and is still a reasonable result below $1/5$ of one degree. According to section 6.1.2.5, the highest value reported is ($\bar{\sigma}_{\bar{\varepsilon}} = 0.160°$, $\bar{\sigma}_{\bar{\varepsilon}_{\%}} \pm 0.956\%$). Therefore, its repeatability is acceptable; even though the accuracy is suffering from some systematic and random errors. Depending on the position (P_i) and the *ERB* band, some of these errors are not tolerable.

8.1 Time domain method

8.1.1 Single reflection scenario

This section describes the discussion of the results of the experiment introduced in section 5.2.1 and the results described in chapter 6. One of the extents found after performing such a comprehensive study of the direct sound and first reflection over 14 positions is the difficulty to explain the random errors and the systematic errors. It could be said that it is frequency and spatially dependent and its accuracy strongly depends on the type of probe chosen. In the quest for minimising errors, several techniques were tested; the repeatability was found not to be a random issue, as most of the positions showed uncertainties below 1 degree. The noticeable outliers had been found in the far positions in both probes. The p-p probe seems to have larger errors and larger uncertainties than the positions mentioned previously. The reason is likely a combination of factors such as the ones mentioned below:

It is important to note that the positions that are far from the source such as $P_{12} - P_{14}$, tend to have smaller differences of time and its impulse response (*IR*) will show that early reflections and the direct sound are nearer each other than in the first positions $P_1 - P_{10}$ and the first reflection (refer to *TOF* results in Appendix B and **Figure 5.9**). These close values on time of flight *TOF* made the analysis of the of the first reflection and the direct sound difficult because they tend to overlap and interfere with each other, and the accuracy of detection of the incoming sound is compromised with low signal-to-noise ratio (*SNR*) and time smearing on the impulse response (*IR*). This may be the reason for poor results in these positions. It seems that even with a poor directivity to capture the first reflection at positions near the source ($P_1 - P_{10}$), it was possible to obtain better angle estimations than in the far positions from the source (e.g. $P_{11} - P_{14}$).

The estimation of the equivalent rectangular bands (*ERB*) shows that the first band gives bad estimations because the band is interfering with the cut-off frequency of the studio

monitor *Genelec* 8030A, which is around 66 Hz, and also because the microphone sensitivity of the p-p probe at low frequencies is not reliable. Therefore, from second band (ERB_2) up to the seventh band (ERB_7) (0.7 - 5 kHz) are consistent for detection of the first reflection. Higher frequencies suffer from directivity deformation in both probes, consequently increasing the error of estimation.

A sound wave front is usually planar by the time it reaches the listener's ear because it has travelled a distance that makes its radius big enough to look more like a plane wave shape than a spherical curved shape. However, if the source and the receiver are sufficiently near, the wave front will show a spherical shape and the spectral content of the wave will add emphasis to the lower frequencies over the high frequencies of the source. This limit is about distances larger than one metre (Begault, 2000). Consequently, the far positions suffer from a lack of power on the high frequencies compared with the low frequencies. On the other hand, a monitor speaker has less efficiency radiating low frequencies because it needs a larger driver, a strong housing and magnet, as well as a powerful amplifier to deliver proper low frequency balance over the mid and high frequencies.

Positions $P_{11} - P_{14}$ suffered from being the last ones measured after performing measurements for more than two of hours due to the long duration of recoding 10 repetitions per position with 10 seconds duration per position. Additionally, the way of measuring the p-p probe also created the first sets of positions from P_1 to P_{14} and then a second set of measurements were made with the 90 degree rotated p-p probe. The way of measuring the positions created an incremental error that was larger in the last positions. The lack of precision in these measurements implies that in the future, the distances that are too far and need a combination of rulers to measure them could be made first in order to diminish the probability of making erroneous measurements.

8.2 Short-time Fourier method

There are many of external factors, which can be detrimental to obtaining good readings of the estimation angles. Firstly, if the extraction of the impulse response exhibits artifacts it will translate those anomalies into inaccuracies of the impulse response on the low and high frequencies (ERB_1, and ERB_{11} - ERB_{14}), where it tends to miscalculate the directions and the levels. The accurate deconvolution of the impulse response (IR) is recommended to implement it in time domain (TD). In the case of this thesis, it was not implemented in that way because the sample frequency used (first was 96 and later 192 kHz) prohibited the process using a computer without enough memory resources and process capability. The creation of the exponential sine swept signal (chirp) for the measurements did not implement the ramp in the *Soundfield* and the p-p probe code of deconvolution. Unfortunately it only was implemented in code designed to be used by the *Microflown* probe, which was not reported in this thesis due to lack of data and time to include it. In order to make better estimations of the impulse response, it needs to be implemented in future measurements. In addition, it is worth to mention that the compensation for the low frequencies was also not implemented in the exponential sine swept signal. It was discovered in the spectrogram of the exponential sine swept signal that it exhibited artifacts that happened at the very end of the signal, which may be the cause of the inaccuracies in high frequencies found both in p-p probe and *Soundfield* microphone cases. Some comparison testing was done using the *RME Fireface UFX* sound card and the *Motu* 896HD, and this artefact was still audible in some of the samples. Nevertheless, because this artifact occurred at a very high frequency almost at the end of the sweep, it did not adversely affect the measurements, which had a narrower bandwidth of about 11.250 kHz. However, the lack of compensation of the low frequencies in the exponential sine sweep recommended by Farina (Farina, 2000) may have created many localisation errors at low frequencies, which can be attributed to the random behaviour found in the results, especially with the p-p probe.

The results of using the use of circular statistics increased the error of angles of arrival in overall circular mean ($\bar{\theta}$) and average of circular standard deviation (σ_o) of first

reflection. The effect is that the biggest error dominates over the smaller ones and the calculation returns larger uncertainties than using linear statistics as it was mistakenly reported at first.

The batteries of the preamps used could be exhausted after 2 hours of measurements and by the time the last positions were measured, more than 2 hours had passed. The effect found when the batteries were emptied was that the sound was distorted and the level of the signal was too low.

The majority of the measurements show low variability with bad precision in some of the positions. The cause of the systematic errors depends on additional external causes, which have influenced the results. It is expected that the larger errors occur in the first reflection measurement using either intensity probe as the overall mean absolute percentage errors, and overall average of standard deviation of *MAPE* results, although the position of the outliers is not consistent in all the different methods applied. Some trends tend to be monotonically increasing the error, while others have a minimum value around *ERB* located towards the middle ($ERB_6 - ERB_{10}$).

There is a general issue found with the *Soundfield* microphone measurements. The problem arises when the error of estimation of the direct sound is larger than the localisation Blur for the frontal side. To minimise these errors it is necessary to have better correction filters customised for each microphone capsule of the *Soundfield* microphone unit. Currently, the *Soundfield* company is not interested in improving the accuracy of this microphone to convert it into a measurement-grade instrument. The device mentioned has been commercialised as a studio grade-recording microphone with applications for surround recording. This may be an opportunity to develop the next generation of *Soundfield* microphones with measurement grade capsules or at least with smaller distances to overcome the high frequency limit, which tends to be below 5 kHz[19],[20],[21].

---

[19] http://www.euphonia.fr/pdf/DPA4.pdf [Online accessed on May 23th 2013]
[20] https://www.kickstarter.com/projects/1569945514/brahma-affordable-ambisonics-microphone [Online accessed on May 23th 2013]
[21] http://www.oktavausa.com/ProductsPages/Ambient4DMic.html [Online accessed on May 23th 2013]

The method of post-processing the data introduces many of potential errors. First, the data is recorded with *DAW* software, which depends in several gains on the 4 channels. The third sound card used, the *RME* Fireface *UFX* sound card, is a good option to overcome the problem of how to calibrate the microphone gains in intermediate values of the gain. When the experiments were previously conducted with the *Focusrite Saffire* 26 i/o and the *Motu* 896HD sound cards, it was not possible to use an intermediate gain on the preamps, only maximum or minimum. The *Motu* 896HD sound card has more options than the *Focusrite Saffire 26 i/o*. One of the desirable inputs for the laboratory equipment is to have at least 4 TRS or TS plug inputs. Another desirable feature is to physically have three input gains such as microphone level (used in measurements because of variable gain with the analogue preamp), line-level (-10 dB) and Balanced level (+4dB). Of course, it is a major problem to try to record good signals with the high mismatch that was found using the measurement-grade microphones and the *Microflown* probe. This was the main reason that this research changed the sound card to the *RME Fireface UFX*, as this was the only way to acquire such data with confidence. However, in the experience of using the 2 sets of *Soundfield* microphones, it is worth noting that the outputs of the *Soundfield* microphones were adequate to record with optimum levels to avoid excessive floor noise because of insufficient gain while recording or clipping because of excessive gain.

The editing of the recorded data includes cut to length of the recorded data and its manual naming using the *DAW* software cannot be automated. This became a very time-consuming task due to the fact that 14 positions with 10 measurements of 4 channels give a total of 560 different high-resolution wav files that needed to be edited and named before using them in the post-processing phase in *Matlab*. This data needed great care in order to be analysed without introducing misplaced data.

The *DAW* software *Logic Pro 9* records audio measurements as wav files @ 24 bits and 192 kHz. However, *Matlab* 2011b can read audio files up to 32 bits, and the internal processing is performed with 32 bits, on single precision numbers and with 64 bits with double precision numbers. The best way to improve the signal-to-noise-ratio (*SNR*) on the recording system is to use any *DAW* that can record at 32 bits native. In this way, the

problem of acquiring the signals at low levels because of the impedance mismatch can be ameliorated. At the end of the completed measurements used in this report, the plan was to change the *DAW* to *Pro Tools* 9 because of this feature and the easy way of editing data for preparation for post-processing. Some preliminary recordings for using the *Microflown* probe were recorded in this configuration and with the *RME Fireface UFX* sound card. For future measurements, another option is to acquire the sound recordings used to extract the impulse response directly into *Matlab* by using an application called *Playrec* from the author Robert Humphrey[22]. The downside is that a custom sound card's driver such as the ASIO or PortAudio or a National Instruments custom acquisition A/D card is needed. One of the limitations found when using several applications together is that they were not to designed to work simultaneously and they need a comprehensive test to validate their compatibility. Further research is needed to overcome these problems; the level of technology that is available during its realisation will always limit a given project.

The next stage was the running and extraction of the impulse response. Zero-padding was used to cut the length of the signals in order to make it easy to deconvolve using high sampling rates. This choice created very long wav files and it was difficult to manage them in the *Matlab* post-processing. The chirp length was chosen to be as long as possible (10 s) in order to have enough time to cover the lower frequency octaves. Within the constraints of the computer power assigned to the project, a compromise was made between the accuracy and the time it took to post-process it.

The platform of post-processing was split between *Excel* and *Matlab* in order to perform the ordering of the matrices and data. This introduced additional possible errors. It was necessary to verify all the steps several times to be sure that the results were consistent. After checking these numerous steps, the conclusion is that an automated system should avoid using many cross platforms among the *DAW , Matlab* and *Excel*. Unfortunately, the measurement system developed here is not practical for the acousticians that are not familiar with the platforms used in this thesis and needs to be made offline.

---

[22] http://www.playrec.co.uk/compiling.php [Online accessed on May 13th 2013]

The *Soundfield* microphone delivers better results on the low frequencies (e.g. $ERB_1 - ERB_5$) than in the high frequencies (e.g. $ERB_{11} - ERB_{14}$). This is because the figure-of-eight patterns are not distorted, while at high frequencies the patterns show strong asymmetries, which are the cause of the systematic errors found.

On the other hand, the p-p probe is not adequate at very low (e.g. ERB_1) and at very high frequencies (e.g. $ERB_{14} - ERB_{16}$).

In some of the positions, there were minimum errors on both probes. It is worth noting that position P_6 showed the best-ranked error with both intensity probes.

According to Blauert (Blauert, 1997b), the human localisation blur is between 1º and 2.5º. The results obtained showed an overall mean error of 2.971º for the p-p probe and 15.895º for the *Soundfied* microphone ST350. These results mean that using the p-p probe is the best option, even though it is not practical in all cases. By using this measurement system it is possible to achieve an angular resolution slightly larger than the human localisation blur within a frequency range from 20 Hz up to 10 kHz. It is possible to improve the estimation of this value by the use of filters carefully customised for each microphone by adapting individual calibration. This result contradicts the results that were reported by Günel (Günel et al., 2007), where the mean error for the *Soundfield* SPS422B found was 7º. The difference is that the angles tested in this study did not cover 360º, whereas Günel's experiment did cover 360º with a step angle of 5º, which is the original request of this study. Therefore, it is encouraging to fully validate these results at 360º and to test the wavelet approach, which may improve the present results.

8.2.1 Measurement in small rooms

The following section deals with the discussion of the experiment introduced in section 5.5 and the results presented in Chapter 7.

The direct sound is clearly detected in both **Figure 7.1** and **Figure 7.2**, which corresponds with the white colour on the diffuse estimate values (ψ) in the diffuse

estimate graph. It can be seen that the mean direction estimate for the direct sound on each frequency band is coherent in all the frequency bandwidth (red rectangles) and tends to deviate slightly at frequencies above 8 kHz because of the limitation of the correction filters of the *Soundfield* microphone. The model is capable of detecting very few full bandwidth reflections (green rectangles) in **Figure 7.2** and depending on the orientation of the source, around 13 ms it can sense only one strong coherent reflection. Because of this limitation, in some respects, it is preferable to use arbitrarily chosen time windows to detect reflections as Ohta et al. had implemented before (Ohta et al., 2008, Yano et al., 2008). The reason for the appearance of some blur information which does not convey any reflection information may be using of overlapping windows with 50 %, which may create some type of redundant information. This choice creates a double average because each window intersects common information and may create some ambiguous direction estimates. The choice of a stepper window rather than the long Hann window is compromised with the distortion, which may be found in its spectrum.

On the other hand, the practical application of Fourier transform by the use of the *Matlab*'s `fft` creates a scaling problem which is partially solved by normalizing its instantaneous complex value dividing it by the length of the signal. However, this approach sometimes needs a careful review before combining the orthogonal components of vectors, which may underestimate or overestimate spectral peaks. One problem found was that increasing the sampling resolution also creates the need to raise the length of the time window in order to maintain the same frequency resolution. This option is not very useful for the measurement of short decay times, on the grounds that a large number of information is irrelevant, since there is mainly background noise after and before the sound decay. The random behaviour of a signal may be filtered by subsequent averaging, but zero padding was found to be a complex procedure when the magnitudes need to be converted to dB, and also tends to deliver a wrong impression of the spectrum.

It is interesting to note that this type of directional analysis is not completely valid for measurements made with only one source-receiver position, because its response varies considerably with different angles and different source positions. One possible reason is the fact that in small rooms, only statistical measurements are meaningful because of the

high variability of results across changes in spatial, temporal and spectral parameters (Geddes and Lee, 2003b).

8.3 Summary

Several causes of inaccuracies and a high level of systematic and random errors have been addressed. In the case of the laboratory experiments specifically, a very comprehensive analysis is obtained in the single reflection scenario. In the case of the measurements of small rooms, a more practical experiment was presented. The level of detailed information found in the laboratory reveals that the random errors are not easy to avoid even with refinements on the method of acquiring data and the post-processing of the information. The results found in this study help to define a workable frequency range for each of the intensity probes from 66 Hz up to 8 kHz using both probes.

Chapter 9 : Conclusions and Further work

9.1 Conclusions

Two methods have been presented to characterise the time of arrival (TOA), strength and direction of arrival (DOA) of early reflections in small rooms: a) Time domain (TD) and b) Short-time Fourier transform ($STFT$) method based on the spatial impulse response rendering ($SIRR$) method proposed by Merimaa (Merimaa, 2006). The reason for using both is because the time domain method (TD) and the frequency domain method ($STFT$) complement each other in their deficiencies.

The estimation of the direction of the arrival of early reflections using the $STFT$ approach in time-frequency domain appears to be a good technique with high accuracy and repeatability in comparison with the time domain (TD) analysis. However, both analyses are complementary. The inclusion of equivalent rectangular band bank (ERB) filter model helps bridge the gap between the objective measures and the subjective measures.

The measurement method applied to two different intensity probes still found inaccurate estimations on positions $P_{11} - P_{14}$. There are several possible reasons for the incorrect estimates in the positions, which are far from the source. The systematic one is caused by the increasing interference between the direct sound and first reflection, as the time of flight of both signals is closer. A careful study of shallow reflections is proposed by using a single point sound source to eliminate the effect of the two drivers on the sound source.

In the case of the *Soundfield* microphone as an affordable 3-D intensity probe, it is possible to be used in practical measurements in small rooms. However, the accuracy of this device is not for high resolution, as angular errors are too high to be useful to be trusted for remedial room acoustic treatment. The physical limitation of the measuring system is the resolution of the *Soundfield* microphone model SP 422, especially in the

forward and backward directions where the human localisation blur angle is more accurate. The future use of the *Microflown* probe may improve the measurement results, and better prepare the impulse response to create more accurate auralisation.

Using these methods, two variants have been compared: a) instantaneous intensity ($\vec{i}_{inst}$), and b) complex instantaneous intensity ($\vec{i}_{complex}$).

The accuracy of the angle of the arrival of direct sound (θ_{s_i}) and first reflection (θ_{m_i}) case had been validated using two measurement probes:

a) The p-p intensity probe with an overall mean error $\bar{\varepsilon}_{\bar{\varepsilon}}$ = 2.971° ±0.414°, which in terms of percentage is 6.866% ±0.251%. This result was found using the instantaneous intensity $\langle \vec{I}_a \rangle$ average on each position method across the 16 *ERB* bands.

b) The *Soundfield* microphone model ST350, which has an overall mean error $\bar{\varepsilon}_{\bar{\varepsilon}}$ =15.895° ±1.248° which in terms of percentage is 24.764% ±1.366%. This result was found using the complex instantaneous intensity ($\vec{i}_{complex}$) method in Time Domain (*TD*).

The rigorous method of treating the angles with circular statistics showed a slight increase in the best case on accuracy. The value of 2.971° seems slightly high; nevertheless, it is worth noting that the outliers found in the first and last *ERB*s were not discharged in light of showing the deficiencies of this measurement method. If they were discharged, a very accurate value could be reported.

The mean absolute percentage error (*MAPE*) is a practical way of showing the errors, which can be compared easily against the different probes as a positive percentage. However, in academic literature, the accuracy of an angular measurement system still needs to be reported as an error in terms of angle. Therefore, a dual error system and the circular statistics are needed, while linear statistics method is necessary for the *MAPE* calculations.

These results are not consistent regarding which method applied yields better results for measuring the accuracy of the measurement system. This may be caused by the difference in the time consumed to acquire a 3-D impulse response. The probe that acquires the 3-D data in one step is more likely to have less random errors than the one that needs to rotate the probe to make another set of measurements in order to obtain the complete 3-D impulse response. This is disadvantage, as it forces a complete analysis on each measurement in order to decide which of the techniques delivers better results.

If the real values of the measurement are not known, then it is important to choose the p-p probe and stick with the lowest uncertainty the device provides. It may be possible to use another intensity probe. The important point is to be able to know the uncertainty of the device prior to using it, otherwise the estimation of the incoming reflections may be biased, and the exact amount it is biased will never be known.

One of the most sensitive parts of the measurement system is the deconvolution of the impulse response algorithm. The important points are that time domain convolution with the inverse of the sine sweep signal is the best solution in terms of not using any fast Fourier transform and inverse Fourier transform in the process, which may introduce rounding errors by the complex numbers involved in the algorithm. In this thesis it was not possible to implement this algorithm due to the long signals acquired with a high sampling rate. It was a compromise between using more detail for the detection of reflections and more accuracy on the extraction of the impulse response at the expense of losing time resolution. Another limitation was to ensure that the beginning and the end of the signal are crossing zero point in order to avoid spectral aliasing when the fast Fourier transform is applied. If this is implemented as a linear time ramp, it creates a temporal linear gain in the impulse response (IR). If the sine sweep signal is created with visual inspection cutting the non-zero crossing at the beginning and end of the sine swept signal, it becomes not practical, although it ensures the quality of the deconvolution process. The pre-emphasis on low frequencies was not implemented in the exponential sine sweep. This can affect the low signal-to-noise ratio, which could bias the angles of arrival at low frequencies, especially in the far positions ($P_{11} - P_{14}$).

The use of *DAW* recording equipment combined with the high impedance of the acoustic measurement-grade p-p probe, is problematic. This is due to the fact that consumer recording equipment and the low gain of the internal preamps of a measurement-grade microphone imposes a high gain in the preamps of the audio interface, and multiple channels with similar calibrated values. The compromise is using a high quality analogue-to-digital conversion and a high sample rate, which is not common in measurement grade-equipment such as the Pulse system from Brüel and Kjaer.

The use of the dampening vibration devices in the microphone stands used for the source and the receiver is highly recommended in order to minimize the random errors found in the low and high frequencies in the directional data. It helped to achieve the results reported in this thesis.

The overall mean error ($\bar{\varepsilon}_{\bar{\varepsilon}}$) was estimated using the *ERB* bands from the *STFT* method. Both time-domain method (*TD*) and short-time Fourier transform method (*STFT*) are complementary.

In the case of this study, a sound source using two drivers (*Genelec* 8030A studio speaker) was selected. This is typically the case encountered in critical listening rooms, where the source is usually a multi-driver monitor speaker. If better results need to be obtained, it is recommended to use of an omni-directional source across the usable frequency range such as the OmniSource type 4295 from Brüel and Kjaer[23], which is based on by design by Pollack (Polack, 1996).

9.2 Further work

- To perform similar angular accuracy measurements using the *Microflown* USP p-u intensity probe to validate the resolution of the probe for detection of sound reflections. The author expects that this will improve the results obtained with the p-p intensity probe so far.

---

[23] http://www.bksv.com/doc/bp1689.pdf [Online accessed on May 25th 2013]

- Further investigation of the proposed use of the quantity called complex instantaneous intensity ($\vec{I}_{complex}$) inside the short-time Fourier transform method ($STFT$) is needed.

- To characterise and use an omni-directional source based on a compressor driver and a tube terminated end to improve the accuracy of detection of reflections. If possible, use the Brüel and Kjaer OmniSource type 4295.

- To study simultaneous arrival of reflections in real rooms using the present technique and combining dual array to resolve ambiguities, possibly using two intensity probes or two *Soundfield* type microphones and a triangulation technique based on the work of Dimoulas (Dimoulas et al., 2007) and the advanced reflection techniques proposed by Tervo (Tervo, 2012).

- Research on the link of perception of early reflections with objective measurements that can rank acoustic quality of small rooms. Develop an objective measure for acoustic quality based on directionality of reflected energy, and design perceptual experiments to validate it according to the current perceptual research in spatial audio.

- Study the case of multiple reflections in 3-D *IR* using laboratory and real rooms.

- Design an automation of the 3-D *IR* measurement system performed with mechanical devices to increase the accuracy and reliability of the measurement system.

9.2.1 Other measurements

The measurements planed were not completely performed with the *Microflown* probe. They need to be validated and compared under the same conditions with the p-p intensity probe and the *Soundfield* microphone model ST 350, in order to report which device exhibits the best accuracy in the market. This is still an open question for research. The method developed by the author is suitable for comparison of different intensity probes.

The acoustic camera may be tested in similar cases to obtain its accuracy on detection of reflected energy in controlled environments. The transient sound phenomenon has been less investigated than the steady state noise sources.

The design and construction of a couple of custom B-Format microphones, each of them comprised of 4 Tascam cardiod 5 mm diameter capsules, may be beneficial in order to minimize the distance between the capsules. The creation of an accurate tetrahedral enclosure made by *CNC* permits a good quality assembly. The pre-amplification of the signals may be done with a pair of Tascam DR100, which sum four inputs and once its accuracy is validated in the semi-anechoic chamber, it is possible to measure more complex sound fields. This project may help to improve the detection of simultaneous arrivals, but at an expense of more channels to analyse following the Dimoulas approach (Dimoulas et al., 2007).

The recent introduction to the market of the *Soundbrush* probe made by collaboration of G.R.A.S. and LMS companies may be tested on its performance on cases such as transient signals where there is sound decay. Some research has been done with intensity probes using a tetrahedral microphone array embedded in a spherical body, although at this time there is no publication on its accuracy on detection of early reflections, only on detection of a broadband source with limitations of errors about 10º at 4 kHz (Janssens et al., 2013). However, it seems to be a good option since the spherical enclosure helps to obtain a uniform response on all directions and the diffraction effect of a sphere can be compensated.

It may be interesting to test the performance of a six-microphone intensity probe such as the Type 50AI from G.R.A.S.[24], which is the most expensive solution, and additionally uses two additional channels, ensuring high accuracy. According to the accuracy results measured with the one-dimensional p-p intensity probe; it may be the most accurate measuring system for detection of early reflections in enclosures.

---

[24] http://www.gras.dk/media/docs/files/items/p/d/pd_50vi_ver_06_08_02.pdf
[Accessed on Oct 21st 2013].

9.2.2 Proposal of an objective metric

By using the current perceptual knowledge of the precedence effect and image-shift, it is possible to predict the perception of detrimental and desirable reflections once they are extracted from a three-dimensional impulse response ($3\text{-D } IR$). The main idea is to compare the measurement of a three-dimensional hedgehog distribution to a set of given ideal room impulse responses (IR), with desirable preferences of temporal and directional reflections in order to assess the quality of a given space at a given coordinate. This work is a multi-dimensional analysis because it requires mixing different independent room parameters such as room geometry, acoustic absorption of the boundaries and fittings in the room. In addition, it is needed to determine multiple listening positions in each enclosure in order to fully characterise them.

9.2.3 Short-time Fourier transform ($STFT$) based

The use of two-dimensional functions such as short-time Fourier transform ($STFT$) may give information for temporal and spectral behaviour related to the quality of the reflections. This is exploiting the strong correlation between a successful objective measure and the perception of a desirable acoustic quality. These experiments will use the tool developed in the present thesis and require an extensive number of measured spaces to make a preference table.

9.2.4 Use diffuseness estimate to stop early reflection analysis

The mixing time is a parameter that can help to determine the boundaries of the early reflections and the late reverberation. Research of proper parameters to map diffuseness estimate (ψ) may find a use for determining the mixing time (t_{mix}).

Appendix A: *Matlab* scripts

There are several scripts used for creating the results of this report.

For the p-p intensity probe the script used to analyse the signals is named:

`Extract_winMLS_signals_Corrected69LOG_BFormat_192kHz.m`

In the case of the *Soundfield* microphone model ST350 the script used to analyse the signals is named:

`Extract_winMLS_signals_Corrected70LOG_BFormat.m`

Each script is too long to attach it to the body of the report, so it was included in the DVD ROM in order to have the complete set of data to analyse and the scripts used.

In the case of post-processing the signals to obtain the desired statistics, there are several variants of scripts used depending on the case studied:

`Postprocessing_reflections_ERBs_odds_Bformat.m`

`Postprocessing_reflections_ERBs_odds_Bformat_New.m`

`Postprocessing_reflections_ERBs_odds_p_p_probe.m`

`Postprocessing_reflections_ERBs_odds_p_p_probe_New.m`

`Postprocessing_reflections_ERBs_odds.m`

`Postprocessing_reflections_TD_odds_BFormat_Complex_intensity_N.m`

`Postprocessing_reflections_TD_odds_BFormat_Inst_int_Avera.m`

`Postprocessing_reflections_TD_odds_BFormat_Inst_int_Avera_N.m`

`Postprocessing_reflections_TD_odds_BFormat_Inst_int_Avera_NC.m`

`Postprocessing_reflections_TD_odds_BFormat_Inst_int_Avera_NCC.m`

`Postprocessing_reflections_TD_odds_BFormat_inst_intensity.m`

```
Postprocessing_reflections_TD_odds_BFormat_inst_intensity_N.m
```

```
Postprocessing_reflections_TD_odds_p_p_probe_Complex_intensity.m
```

```
Postprocessing_reflections_TD_odds_p_p_probe_Complex_intensity_N.m
```

```
Postprocessing_reflections_TD_odds_p_p_probe_inst_intensity.m
```

```
Postprocessing_reflections_TD_odds_p_p_probe_intst_Avera.m
```

```
Postprocessing_reflections_TD_odds_p_p_probe_intst_Avera_N.m
```

```
Postprocessing_reflections_Time_Domain_odds_p_p_probe.m
```

```
Postprocessing_reflections_TD_odds_p_p_probe_intst_Avera_rel_error.m
```

Appendix B: Calculation method of the single reflection scenario

This is calculation is performed in the *Matlab* script named

a) `Extract_winMLS_signals_Corrected69LOG_BFormat_192kHz.m`

b) `Extract_winMLS_signals_Corrected70LOG_BFormat.m`

The angle of arrival of first reflection (θ_{m_i}) is selected arbitrarily by the angular step used and its calculation was made with the geometry presented in **Figure 5.8**.

The calculation of the angle of arrival of the first reflection (θ_{m_i}) using equation (4.1) in page 63 yields the following numbers:

`theta_m1  = -75.067°`

`theta_m2  = -74.067°`

`theta_m3  = -73.067°`

`theta_m4  = -72.067°`

`theta_m5  = -71.067°`

`theta_m6  = -70.067°`

`theta_m7  = -68.067°`

`theta_m8  = -66.067°`

`theta_m9  = -64.067°`

`theta_m10 = -62.067°`

`theta_m11 = -60.067°`

`theta_m12 = -55.067°`

theta_m13 = -50.067°

theta_m14 = -45.067°

The calculated coordinates in distances in equation (4.2) in page 63 gives the following data in meters:

x_1   = 0.707 m

x_2   = 0.758 m

x_3   = 0.807 m

x_4   = 0.858 m

x_5   = 0.909 m

x_6   = 0.961 m

x_7   = 1.068 m

x_8   = 1.177 m

x_9   = 1.289 m

x_10 = 1.406 m

x_11 = 1.526 m

x_12 = 1.852 m

x_13 = 2.219 m

x_14 = 2.645 m

The increments of distance Δx_i along x-axis for the generation of the step angles are calculated with equation (4.3) in page 64:

delta_x_2 = 0.0498 m

delta_x_3 = 0.0503 m

```
delta_x_4  = 0.0508 m
```

```
delta_x_5  = 0.0514 m
```

```
delta_x_6  = 0.0520 m
```

```
delta_x_7  = 0.1061 m
```

```
delta_x_8  = 0.1091 m
```

```
delta_x_9  = 0.1126 m
```

```
delta_x_10 = 0.1164 m
```

```
delta_x_11 = 0.1208 m
```

```
delta_x_12 = 0.3252 m
```

```
delta_x_13 = 0.3675 m
```

```
delta_x_14 = 0.4256 m
```

The computation of the angles of arrival of the direct sound (θ_{s_i}) use the *x*-axis coordinate values by using the equation (4.4) in page 64:

```
theta_s_1  = 45.000°
```

```
theta_s_2  = 43.052°
```

```
theta_s_3  = 41.218°
```

```
theta_s_4  = 39.491°
```

```
theta_s_5  = 37.865°
```

```
theta_s_6  = 36.331°
```

```
theta_s_7  = 33.517°
```

```
theta_s_8  = 31.002°
```

```
theta_s_9  = 28.742°
```

```
theta_s_10 = 26.702°
```

```
theta_s_11 = 24.853°
```

```
theta_s_12 = 20.898°
```

```
theta_s_13 = 17.671°
```

```
theta_s_14 = 14.966°
```

The distance travelled of the direct sound from the source to the receiver ($r_{i_{direct}}$) is calculated with equation (4.5) in page 64.

The time delay (t_i) of the ith direct sound is calculated using equation (4.6) in page 65, applying it for the 14 positions it results in:

```
Time_delay_direct_sound_1  = 2.937 ms
```

```
Time_delay_direct_sound_2  = 3.043 ms
```

```
Time_delay_direct_sound_3  = 3.152 ms
```

```
Time_delay_direct_sound_4  = 3.266 ms
```

```
Time_delay_direct_sound_5  = 3.384 ms
```

```
Time_delay_direct_sound_6  = 3.506 ms
```

```
Time_delay_direct_sound_7  = 3.761 ms
```

```
Time_delay_direct_sound_8  = 4.033 ms
```

```
Time_delay_direct_sound_9  = 4.319 ms
```

```
Time_delay_direct_sound_10 = 4.622 ms
```

```
Time_delay_direct_sound_11 = 4.942 ms
```

```
Time_delay_direct_sound_12 = 5.823 ms
```

```
Time_delay_direct_sound_13 = 6.842 ms
```

```
Time_delay_direct_sound_14 = 8.043 ms
```

The time difference also known as time of flight (TOF) from the first reflection to the direct sound is calculated as the measure to assess the reliability of the impulse responses using equation (4.7) and its expanded version (4.8) in page 65. The time delay calculated in milliseconds yield the following results:

```
difference_direct_first_reflection_1  = 5.123 ms
```

```
difference_direct_first_reflection_2  = 5.057 ms
```

```
difference_direct_first_reflection_3  = 4.989 ms
```

```
difference_direct_first_reflection_4  = 4.920 ms
```

```
difference_direct_first_reflection_5  = 4.850 ms
```

```
difference_direct_first_reflection_6  = 4.779 ms
```

```
difference_direct_first_reflection_7  = 4.634 ms
```

```
difference_direct_first_reflection_8  = 4.488 ms
```

```
difference_direct_first_reflection_9  = 4.341 ms
```

```
difference_direct_first_reflection_10 = 4.193 ms
```

```
difference_direct_first_reflection_11 = 4.045 ms
```

```
difference_direct_first_reflection_12 = 3.677 ms
```

```
difference_direct_first_reflection_13 = 3.314 ms
```

```
difference_direct_first_reflection_14 = 2.958 ms
```

The difference of adjacent position's time delay ($\Delta t_{i \to i+n}$) is calculated using equation (4.9) in page 66:

```
difference_of_Time_delay_1_vs_2   = 0.066 ms

difference_of_Time_delay_2_vs_3   = 0.068 ms

difference_of_Time_delay_3_vs_4   = 0.069 ms

difference_of_Time_delay_4_vs_5   = 0.070 ms

difference_of_Time_delay_5_vs_6   = 0.071 ms

difference_of_Time_delay_6_vs_7   = 0.144 ms

difference_of_Time_delay_7_vs_8   = 0.146 ms

difference_of_Time_delay_8_vs_9   = 0.147 ms

difference_of_Time_delay_9_vs_10  = 0.1479 ms

difference_of_Time_delay_10_vs_11 = 0.148 ms

difference_of_Time_delay_11_vs_12 = 0.368 ms

difference_of_Time_delay_12_vs_13 = 0.363 ms

difference_of_Time_delay_13_vs_14 = 0.356 ms
```

List of references

(1994). British Standards, BS EN 61043:1994 IEC 1043:1993 Electroacoustics — Instruments for the measurement of sound intensity — Measurement with pairs of pressure sensing microphones.

(2002). The 3-Dimensional Acoustic Probe 3-DP. Meyer Sound Laboratories. Catalogue: Available from:

http://www.microflown.com/files/media/library/Applicationnotes/meyersound_3dp.pdf [Accessed on Dec. 10th, 2012].

(2005a). Vector Intensity Probe Type 50VI. G.R.A.S. 3D PROBE. Catalogue: Available from:

http://www.gras.dk/media/docs/files/items/p/d/pd_50vi_ver_06_08_02.pdf [Accessed on Oct 21th 2013].

(2005b). X-Y-Z Sound-intensity Probe Type 50VX G.R.A.S. 3D PROBE. Catalogue: Available from:

http://www.gras.dk/50vx-intensity-probe.html [Accessed on May 4th 2013].

(2007). Data Sheet Genelec 8030A Active Monitoring System. Genelec Oy, Finland. Catalogue: Available from:

http://www.genelec.com/documents/datasheets/DS8030a.pdf [Accessed on April 25th 2013].

(2008). British Standards, BS EN ISO 3382-2:2008, Acoustics — Measurement of Room Acoustic Parameters — Part 2: Reverberation Time in Ordinary Rooms.

(2012). Sound Sources and Impact Sound Source for Building Acoustics: Sound Sources: OmniSourceTM Type 4295 and OmniPowerTM Type 4292-L Impact Sound Source: Tapping Machine Type 3207 Power Amplifiers Type 2734-A and 2734-B. Brüel and Kjaer. Catalogue: Available from:

http://www.bksv.com/doc/bp1689.pdf [Accessed on Sep. 13th, 2012].

Abdou, A. a. M. (1994). *Transient Sound Intensity Measurement for Evaluating the Spatial Information of Sound Fields in Reverberant Enclosures*. Ph.D. Thesis, Concordia University, Montreal, Quebec, Canada, pp. 297.

Abramowitz, M. & Stegun, I. A. (1972). *Handbook of Mathematical Function with Formulas, Graphs, and Mathematical Tables, Tenth Printing*, National Bureau of Standards, Applied Mathematics Series, pp. 1060, 55, p. 14, 948, 990.

Ahonen, J. & Pulkki, V. (2009). Diffuseness Estimation using Temporal Variation of Intensity Vectors. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics.

Ahonen, J., Pulkki, V., Kuech, F., Kallinger, M. & Schultz-Amling, R. (2008). Directional Analysis of Sound Field with Linear Microphone Array and Applications in Sound Reproduction. AES 124th Convention, May 17–20, Amsterdam, The Netherlands.

Alfredson, R. (1980). The Direct Measurement of Acoustic Energy in Transient Sound Fields. *Journal of Sound and Vibration,* **70**(2), p. 181-186.

Ando, Y. & Alrutz, H. (1984). Perception of Coloration in Sound Fields in Relation to the Autocorrelation Function. *Journal of the Acoustical Society of America,* **71**(3), p. 616-618.

Ballou, G. (1991). *Handbook for Sound Engineers - The New Audio Cyclopedia, Third edition,* U.S.A., Howard W Sams and Co, pp. 1525, p. 149, 675, 1146-1147.

Barron, M. (1971). The Subjective Effects of First Reflections in Concert Halls-The Need for Lateral Reflections. *Journal of Sound and Vibration,* **15**(4), p. 475-494.

Barron, M. (1973). Growth And Decay of Sound Intensity in Rooms According to Some Formulae of Geometric Acoustics Theory. *Journal of Sound and Vibration,* **27**(2), p. 183-196.

Barron, M. (2009). *Auditorium Acoustics and Architectural Design, Second,* Oxon, U.K., Spon Press an Imprint of Taylor & Francis, pp. 504, p. 16, 54, 68–69,82, 84, 90, 91, 98–101, 109, 226.

Barron, M. & Marshall, A. H. (1981). Spatial Impression due to Early Lateral Reflections in Concert Halls: the Derivation of a Physical Measure. *Journal of Sound and Vibration,* **77**(2), p. 211-232.

Bassuet, A. (2010). New Acoustical Parameters and Visualization Techniques to Analyze the Spatial Distribution of Sound in Music Spaces. Proceedings of the International Symposium on Room Acoustics, ISRA, August 29-31, Melbourne, Australia.

Basten, T. G. H., De Bree, H. E. & Yntema, D. R. (2009). An Orientation Calibration Procedure for two Acoustic Vector Sensor Configurations. ICSV16th International Congress on Sound and Vibration, July 5-7, Kraków, Poland.

Bech, S. (1995). Timbral aspects of reproduced sound in small rooms. I. *Journal of the Acoustical Society of America,* **97**(3), p. 1717-1726.

Bech, S. (1996). Timbral Aspects of Reproduced Sound in Small Rooms. II. *Journal of the Acoustical Society of America,* **99**(6), p. 3559-3549.

Begault, D. R. (2000). *3-D Sound for Virtual Reality and Multimedia NASA/TM—2000–000000,* U.S.A., National Aeronautics and Space Administration, Ames Research Center, Moffett Field, California 94035, pp. 1-246, p.

Begault, D. R., Mcclain, B. U. & Anderson, M. (2004). Early reflection thresholds for anechoic and reverberant stimuli within a 3-D sound display. Proceedings of the 18th International Congress on Acoustics, Kyoto, Japan.

Begault, D. R., Mcclain, B. U. & Anderson, M. R. (2001). Early reflection thresholds for virtual sound sources. Proceedings of the 2001 International Workshop on Spatial Media, Oct. 25-26, Aizu-Wakamatsu, Japan.

Benjamin, E. (2010). A *Soundfield* Microphone using Tangential Capsules. AES 129th Convention November 4-7, San Francisco, CA, U.S.A.

Beranek, L. L. (2004). *Concert Halls and Opera Houses,* New York U.S.A., Springer-Verlag, pp. 661, p. 524-526.

Beranek, L. L. (2008). Concert Hall Acoustics—2008*. *Journal of the Audio Engineering Society,* **56**(7/8), p. 532-544.

Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software,* **31**(10), p. 1-21.

Billauer, E. (2011). *Peakdet: Peak detection using MATLAB* [Online]. Available: http://www.billauer.co.il/peakdet.html [Accessed on 29 April 2011].

Binelli, M., Venturi, A., Amendola, A. & Farina, A. (2011). Experimental Analysis of Spatial Properties of the Sound Field Inside a Car Employing a Spherical Microphone Array. AES 130th Convention May 13-16, London, U.K., Apr 10.

Blauert, J. (1997a). *Spatial Hearing: The Psychophysics Of Human Sound Localization (Revised Ed'97),* Cambridge, Massachusetts, U.S.A., pp. 495, p. 37-50.

Blauert, J. (1997b). *Spatial Hearing: The Psychophysics of Human Sound Localization (revised edition 97),* Cambridge Massachussetts, MIT Press, pp. 495, p. 14, 41, 44, 49, 222, 276.

Blumlein, A. D. (1933). Improvements in and Relating to Sound-Transmission, Sound-Recording and Sound-Reproducing Systems. British Patent Office. patent number No. 34,657/31. United Kingdom.

Bolt, R. & Roop, R. (1950). Frequency Response Fluctuations in Rooms*. *Journal of the Acoustical Society of America,* **22**(2), p. 280-289.

Bolt, R. H. (1939). Normal Modes of Vibration in Room Acoustics: Angular Distribution Theory. *Journal of the Acoustical Society of America,* **11** p. 74-79.

Bonsi, D., Cengarle, G. & Stanzial, D. (2007). Intensimetric Monitoring of Acoustic Quadraphonic Recordings Reproduced Through a 5.1 Loudspeaker Array. 19th International Congress on Acoustics, 2-7 September, Madrid, Spain., Jun 11.

Burns, T. H. (1996). Measurement and Visualization of Instantaneous Power Flow in Steady-State Acoustic Fields. *Journal of the Acoustical Society of America,* **99**(4), p. 2444-2445.

Cavanaugh, W. J., Tocci, G. C. & Wilkes, J. A. (2009). *Architectural Acoustics,* N.J., U.S.A., John Wiley & Sons, pp. 352, p. 211.

Cazzolato, B. (1999). *Sensing Systems for Active Control of Sound Transmission into Cavities.* Ph.D. Thesis, The University of Adelaide,, South Australia, Australia, pp. 451.

Cazzolato, B., Halim, D., Petersen, D., Kahana, Y. & Kots, A. (2005). An Optical 3D Sound Intensity and Energy Density Probe. Proc Acoustics, November 9-11, Busselton, Western Australia.

Cengarle, G. (2012). *3D audio technologies: applications to sound capture, post-production and listener perception.* Ph.D. Thesis, Pompeu Fabra University, Barcelona, Catalunya, Spain, pp. 167.

Choi, C., Kim, L.-H., Doo, S., Oh, Y. & Sung, K.-M. (2001). Research on the Improvement of Naturalness in Auralization. AES 111th Convention, September 21–24, New York, N.Y., U.S.A., Aug 09.

Choi, C., Kim, L.-H., Oh, Y., Doo, S. & Sung, K.-M. (2003). Measurement of Early Reflections in a Room with Five Microphone System. *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences,* **E86-A**(12), p. 3283-3287.

Chung, J. Y. (1978). Cross−spectral method of measuring acoustic intensity without error caused by instrument phase mismatch. *Journal of the Acoustical Society of America,* **64** p. 1613-1616.

Cotterell, P. S. (2009). *On the Theory of the Second-Order Soundfield Microphone.* Ph.D. Thesis, University of Reading, United Kingdom, pp. 188.

Cox, T. J., Davies, W. J. & Lam, Y. W. (1993). The Sensitivity of Listeners to Early Sound Field Changes in Auditoria. *Acta Acustica united with Acustica* **79**(1), p. 27-41.

Craven, P. & Gerzon, M. A. (1977). Coincident Microphone Simulation Covering Three Dimensional Space and Yielding Various Directional Outputs. Patent patent application 593,244. patent number 4,042,779. U.S.A.

Craven, P., Law, M. & Travis, C. (2010). Microphone Array. Patent, patent application 12/444,628. patent number US 2010/0142732 A1 United States.

D'antonio, P. & Eger, D. (1986). T/60-How Do I Measure Thee, Let Me Count the Ways. AES 81st Convention, November 12-16, Los Angeles, CA, U.S.A.

Da Silva Vieira De Melo, G. (2002). *Measurement and Prediction of Sound Absorption of Room Surfaces and Contents at Low Frequencies.* Ph.D. Thesis, Universidade Federal De Santa Catarina, Florianópolis, Janeiro, Brasil, pp. 197.

Damaske, P. & Ando, Y. (1972). Interaural Crosscorrelation for Multichannel Loudspeaker Reproduction. *Acta Acoustica united with Acustica,* **27**(4), p. 232-238.

Daniel, J. R. M. (2003). Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format. AES 23rd International Conference on Signal Processing in Audio Recording and Reproduction.

Davis, D. & Patronis, E. (2006). *Sound System Engineering, Third edition,* U.S.A., Focal Press, pp. 489, p. 131-151, 173, 175-188.

De Bree, H. E. (1997). *The Microflown Second Edition.* Ph.D. Thesis, University of Twente, Enschede, The Netherlands, pp. 187.

De Bree, H. E. (2003). An Overview of Microflown Technologies. *Acta Acustica United with Acustica,* **89**(1), p. 163-172.

De Bree, H. E., Druyvesteyn, W. F., Berenschot, E. & Elwenspoek, M. (1999). Three-Dimensional Sound Intensity Measurements using Microflown Particle Velocity Sensors. IEEE International Conference on Micro Electro Mechanical Systems, 1999. MEMS '99. Twelfth.

Defrance, G., Daudet, L. & Polack, J. D. (2009). Using Matching Pursuit for Estimating Mixing Time Within Room Impulse Responses. *Acta Acustica united with Acustica* **95** p. 1071-1081.

Del Galdo, G., Taseska, M., Thiergart, O., Ahonen, J. & Pulkki, V. (2012). The Diffuse Sound Field in Energetic Analysis. *The Journal of the Acoustical Society of America,* **131**(3), p. 2141.

Dimoulas, C., Avdelidis, K., Kalliris, G. & Papanikolaou, G. (2007). Sound Source Localization and B-format Enhancement using *Soundfield* Microphone Sets. AES 122nd Convention, May 5-8, Vienna, Austria. , May 11.

Dimoulas, C., Kalliris, G., Avdelidis, K. & Papanikolaou, G. (2009). Improved Localization of Sound Sources Using Multi-Band Processing of Ambisonic Components. AES 126th Convention, May 7–10, Munich, Germany., Apr 03.

Druyvesteyn, W. E. & De Bree, H. E. (1998). A New Sound Intensity Probe; Comparison to the Bruel & Kjaer p-p probe. *AES 104th Convention 1998 May 16-19 Amsterdam, The Netherlands,* p. 1-12.

Elko, G. W. (1984). *Frequency Domain Estimation of the Complex Acoustic Intensity and Acoustic Energy Density.* Ph.D. Thesis, The Pennsylvania State University, Pennsylvania, U.S.A., pp. 376.

Elko, G. W. (2000). Steerable and variable first-order differential microphone array. US Patent, patent application 08/832,553. patent number 6,041,127. U.S.A.

Elko, G. W. & Pong, A. (1997). A Steerable and Variable First-Order Differential Microphone Array. *International Conference on Acoustics, Speech, and Signal Processing, ICASSP-97, IEEE* p. 223-226 vol. 1.

Enroth, S. (2007). *Spatial Impulse Response Rendering of Digital Waveguide Mesh Room Acoustic Simulations.* M.Sc. Thesis, KTH University, Stockholm, Sweden, pp. 48.

Essert, R. (1997). Progress in Concert Hall Design – Developing an Awareness of Spatial Sound and Learning how to Control it. *EBU Technical Review* (Winter), p. 31-39.

Fahy, F. J. (1977). Measurement of Acoustic Intensity Using the Cross−Spectral Density of two Microphone Signals. *Journal of the Acoustical Society of America,* **62**(4), p. 1057-1059.

Fahy, F. J. (1995). *Sound Intensity, Second edition,* London, U.K., E & FN Spon, pp. 316, p. 2, 9, 19, 74, 85, 170, 238.

Fahy, F. J. (2005). *Foundations of Engineering Acoustics,* London U.K., Elsevier Academic Press, pp. 465, p. 74.

Faller, C. & Kolundzija, M. (2009). Design and Limitations of Non-coincidence Correction Filters for *Soundfield* Microphones. AES 126th Convention 2009 May 7–10 Munich, Germany.

Farina, A. (2000). Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. AES 108th Convention, Feb 19-22, Paris, France.

Farina, A. (2007). Advancements in Impulse Response Measurements by Sine Sweeps. AES 122nd Convention, May 5–8, Vienna, Austria.

Farina, A., Amendola, A., Capra, A. & Varani, C. (2011). Spatial Analysis of Room Impulse Responses Captured with a 32-Capsules Microphone Array. AES 130th Convention May 13-16, London, U.K., Apr 10.

Farina, A. & Ayalon, R. (2003). Recording Concert Hall Acoustics for Posterity. AES 24th International Conference: Multichannel Audio, The New Reality, 26-28 June, Banff, Canada. .

Farina, A. & Ugolotti, E. (1998). Software Implementation of B-Format Encoding and Decoding. AES 104th Convention, May 16-19, Amsterdam, The Netherlands.

Fastl, H. & Zwicker, E. (2007). *Psychoacoustics Facts and Models, Third edition,* New York, U.S.A., Springer-Verlag Berlin Heidelberg, pp. 470, p. 186, 196, 235, 390.

Fazenda, B. M. & Romero-Perez, J. (2008). 3-Dimensional Room Impulse Response Measurements in Critical Listening Spaces. Reproduced Sound 24: Immersive Audio, Apr 30, Brighton U.K. Proceedings of the Institute of Acoustics, **30** (6).

Ferguson, B. G. & Quinn, B. G. (1994). Application of The Short-Time Fourier Transform and the Wigner-Ville Distribution to the Acoustic Localization of Aircraft. *Journal of the Acoustical Society of America,* **96** p. 821-821.

Fisher, N. I. (1971). *Comment on "A Method for Estimating the Standard Deviation of Wind Directions"* [Online]. Available: http://journals.ametsoc.org/doi/pdf/10.1175/1520-0450(1983)022%3C1971%3ACOMFET%3E2.0.CO%3B2%5D.

Flandrin, P., Auger, F. & Chassande-Mottin, E. (2003). Time-Frequency Reassignment— From Principles to Algorithms. *Applications in Time-Frequency Signal Processing.* A Papandreu-Suppappola, p. 179-203.

Friedrich, R. (2012). Sound field microphone. European Patent Office, patent application 05450110.1. patent number EP 1 737 268 B1. Germany.

Fukushima, Y., Suzuki, H. & Omoto, A. (2006). Visualization of Reflected Sound in Enclosed Space by Sound Intensity Measurement. *Acoust. Sci. & Tech.,* **27**(3), p. 1-3.

Gade, S. (1982a). Sound Intensity (Part 1. Theory), Brüel and Kjaer, Technical Review, Volume 3, Nærum, Denmark.

Gade, S. (1982b). Sound Intensity (Part II. Instrumentation & Applications); Flutter Compensation of Tape Recorded Signals for Narrow Band Analysis, Brüel and Kjaer, Technical Review, Volume 4, Nærum, Denmark.

Gade, S. (1985). Validity of Intensity Measurements in Partially Diffuse Sound Field, Brüel and Kjaer, Technical Review, Volume 4, Nærum, Denmark.

Geddes, E. R. (1998). Small Room Acoustics in the Statistical Region. AES Proceedings of the 15th International Conference.

Geddes, E. R. (2009). Audio Acoustics in Small Rooms Presentation Slides, GedLee LLC, [Available from: http://www.gedlee.com Access 2009].

Geddes, E. R. & Lee, L. (2003a). *"Premium Home Theater: Design and Construction", Chapter 2 Hearing and Psychoacoustics* [Online]. Available: http://www.gedlee.com/downloads/Chapter 2.pdf [Accessed on Nov 21th 2009].

Geddes, E. R. & Lee, L. (2003b). *"Premium Home Theater: Design and Construction", Chapter 4 Room Acoustics* [Online]. Available: http://www.gedlee.com/downloads/Chapter 4.pdf [Accessed on Nov 21th 2009].

Gentner, K., Braasch, J. & Calamia, P. (2007). A Perforated Desk Surface to Diminish Coloration in Desktop Audio-Production Environments. AES 123rd Convention, New York, NY, USA, 2007 October 5–8. Audio Engineering Society,

Gerzon, M. A. (1975a). The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound. AES 50th Conference, March.

Gerzon, M. A. (1975b). Recording concert hall acoustics for posterity. *Journal of the Audio Engineering Society,* **23**(7), p. 569–571.

Gover, B. N. (2001). *Development and Use of a Spherical Microphone Array for Measurement of Spatial Properties of Reverberant Sound Fields.* Ph.D. Thesis, University of Waterloo, Waterloo, Ontario, Canada, pp. 388.

Gover, B. N., Ryan, J. G. & Stinson, M. R. (2002). Microphone Array Measurement System for Analysis of Directional and Spatial Variations of Sound Fields. *Journal of the Acoustical Society of America,* **112**(5), p. 1980-1991.

Gover, B. N., Ryan, J. G. & Stinson, M. R. (2004). Measurements of Directional Properties of Reverberant Sound Fields in Rooms Using a Spherical Microphone Array. *Journal of the Acoustical Society of America,* **116**(4), p. 2138-2148.

Griesinger, D. H. (2009). The Importance of The Direct to Reverberant Ratio in The Perception of Distance, Localization, Clarity, and Envelopment. AES 126th Convention, May 7–10, Munich, Germany.

Günel, B., Hacıhabiboğlu, H. & Kondoz, A. (2007). Wavelet Packet Based Analysis of Sound Fields in Rooms using Coincident Microphone Arrays. *Applied Acoustics,* **68**(7), p. 778-796.

Günel, B., Hüseyin, H. & Kondoz, A. M. (2005). Sound Field Analysis using a Compact Pressure-Gradient Array. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, October 16-19, New Paltz, N.Y., U.S.A.

Haas, H. (1972). The Influence of a Single Echo on the Audibility of Speech*. *Journal of the Audio Engineering Society,* **20**(2), p. 146-159. Originally from: H. Haas (1951) "Vber den Einfluss eines Einfachechos auf die Horsamkeit von Sprache", B. R. S. L. C. 363 (1949) and Acustica Vol. 1, Nr. 2, p. 49.

Halmrast, T. (2000). Orchestral Timbre: Comb-Filter Coloration from Reflections. *Journal of Sound and Vibration,* **232**(1), p. 53-69.

Hamilton, J. D. (1994). *Time Series Analysis*, Taylor & Francis US, pp. 799, p.

Hanselman, D. C. & Littlefield, B. (2005). *Mastering MATLAB 7,* New Jersey, U.S.A., Pearson Prentice Hall, pp. 852, p. 371.

Havelock, D., Kuwano, S. & Vorländer, M. (2008). *Handbook of Signal Processing in Acoustics,* N.Y. U.S.A., pp. 1-1900, Volume 2, p. 65-85, 1408-1413, 1421.

Heller, A. & Benjamin, E. M. (2012). Calibration of *Soundfield* Microphones Using the Diffuse-Field Response. AES 133rd Convention, October 26–29, San Francisco, CA, U.S.A.

Heyser, R. C. (1967). Acoustical Measurements by Time Delay Spectrometry. *Journal of the Audio Engineering Society,* **15**(4), p. 371-382.

Heyser, R. C. (1986a). Instantaneous intensity. AES 81st Convention Nov 12-16, Los Angeles, CA, U.S.A.

Heyser, R. C. (1986b). Specific Acoustic Wave Admittance. AES 81st Convention, Nov 12-16, Los Angeles, CA, U.S.A.

Hoeg, W., Christensen, L. & Walker, R. (1997). Subjective Assessment of Audio Quality – The Means and Methods within the EBU. *EBU Technical Review Winter 1997,* (Winter), p. 40-50.

Hopkins, C. (2007). *Sound Insulation, First edition,* Oxford, U.K., Elsevier LTD, pp. 649, p. 48, 54, 97, 480.

Howard, D. M. & Angus, J. A. (2001). *Acoustics and Psychoacoustics, Second edition,* Oxford, U.K., Focal Press an Imprint of Elsevier, pp. 385, p. 97-104.

Huang, Y. & Benesty, J. E. (2004). *Audio Signal Processing for Next-Generation Multimedia Communication Systems,* Boston, Massachusetts, U.S.A., Kluwer Academic Publishers, pp. 389, p. 67-89.

Humphrey, R. (2006). *Automatic Loudspeaker Location Detection for use in Ambisonic Systems*. MEng. Thesis, University of York and KTH, York, United Kingdom, pp. 198.

Hyde, J. R. (1998). Sound Strength in Concert Halls: The Role of The Early Sound Field on Objective and Subjective Measures. *The Journal of the Acoustical Society of America,* **103**(5), p. 2748.

(2008). JCGM 100:2008 GUM 1995 with minor Corrections Evaluation of Measurement Data — Guide to the Expression of Uncertainty in Measurement.

Jacobsen, F. (1979). The Diffuse Field Statistical Considerations Concerning the Reverberant Field in the Steady State, DTU Denmark.

Jacobsen, F. (1991). A Note on Instantaneous and Time-Averaged Active and Reactive Sound Intensity. *Journal of Sound and Vibration,* **47**(3), p. 489-496.

Jacobsen, F. (2002). A Note on Finite Difference Estimation of Acoustic Particle Velocity. *Journal of Sound and Vibration,* **256**(5), p. 849-859.

Jacobsen, F., Cutanda, V. & Juhl, P. M. (1996). A Sound Intensity Probe for Measuring from 50Hz to 10kHz, Brüel and Kjaer Technical Review, Volume 1,

Jacobsen, F. & De Bree, H. E. (2005). Measurement of Sound Intensity: p-u Probes versus p-p Probes. Proceedings of NOVEM, Feb 14, Brazil.

Jacobsen, F. & Nielsen, T. (1987). Spatial Correlation And Coherence In A Reverberant Sound Field. *Journal of Sound and Vibration,* **118**(1), p. 175-180.

Janssens, K., De Weer, D., Bianciardi, F. & Søndergaard, T. (2013). On-Line Sound Brush Measurement Technique for 3D Noise Emission Studies. Safety, Jun Warrendale, PA. Proceedings 13NVC-0192 of the SAE Noise and Vibration Conference, May 2013,

Johansson, M. (2008). *The Hilbert Transform.* M.Sc. Thesis, Växjäo University, Sweden, pp. 36.

Kite, T. (2004). Measurements of Audio Equipment with Log-Swept Sine Chirps. AES 117th Convention, October 28–31, San Francisco, CA, U.S.A.

Kuttruff, H. (1998). Sound fields in Small Rooms. AES 16th international conference: Audio, Acoustics & Amp; Small Spaces.

Kuttruff, H. (2000). *Room Acoustics, Fourth edition,* Oxon, U.K., Spon Press, pp. 350, p. 39, 61, 76, 79, 96-99, 111-115, 248-249, 287.

Kuttruff, H. (2009). *Room Acoustics, Fifth Edition*, pp. 1-389, p. 275-276.

Leis, J. (2003). *Digital Signal Proceesing a Matlab-Based Tutorial Approach,* Baldock, Hertfordshire, U.K., Research Studies Press, pp. 218, p. 109-111.

Li, K. (1995). *Predicting Sound Propagation in Fitted Workrooms.* Master on Applied Science Thesis, The University of British Columbia, Vancouver, Canada, pp. 137.

Lyons, R. (2002). *Understanding Digital Signal Processing* U.S.A., Prentice Hall PTR, pp. 269, p. 89, 406.

Madsen, E. R. (1970). Extraction of Ambiance Information from Ordinary Recordings. *Journal of the Audio Engineering Society,* **18**(5), p. 490-496.

Mäkivirta, A. & Anet, C. (2001). A Survey Study of Today's Monitoring Conditions. Studio Sound, 43(9), p. 1-9.

Merimaa, J. (2006). *Analysis, Synthesis, and Perception of Spatial Sound - Binaural Localization Modeling and Multichannel Loudspeaker Reproduction.* Ph.D. Thesis, Helsinski University of Technology, Espoo, Finland, pp. 196.

Merimaa, J., Lokki, T., Peltonen, T. & Karjalainen, M. (2001). Measurement, Analysis, and Visualization of Directional Room Responses. AES 111th Convention, September 21–24, New York, N.Y., U.S.A. , Jun 19.

Merimaa, J. & Pulkki, V. (2005). Spatial impulse Response Rendering I: Analysis and Synthesis. *Journal of the Audio Engineering Society,* **53**(12), p. 1115-1127.

Meyer, J. & Elko, G. W. (2002). A highly Scalable Spherical Microphone Array Based on an Orthonormal Decomposition of the Soundfield. Proceedings of ICASSP. IEEE,

Meyer, J. & Elko, G. W. (2010). Analysis of the High-Frequency Extension for Spherical Eigenbeamforming Microphone Arrays. *The Journal of the Acoustical Society of America,* **127**(3), p. 1979.

Miah, K. H. (2009). *Three-Dimensional Broadband Intensity Probe for Measuring Acoustical Parameters.* Ph.D. Thesis, The University of Texas at Austin, Austin, Texas, U.S.A., pp. 133.

Miah, K. H. & Hixon, E. (2010). Design and Performance Evaluation of a Broadband Three Dimensional Acoustic Intensity Measuring System. *Journal of the Acoustical Society of America,* **127**(4), p. 2338-2346.

Mills, A. W. (1958). On the Minimum Audible Angle. *Journal of the Acoustical Society of America,* **30**(4), p. 237-246.

Miura, T., Muraoka, T. & Ifukube, T. (2010). Point-Source Loudspeaker Reversely-Attached Acoustic Horn: Improvement of Acoustic Characteristics and Application to some Measurements. AES 129th Convention November 4-7 San Francisco CA, U.S.A.

Moore, B. C. J. (2003). *An Introduction to the Psychology of Hearing,* London, U.K., Elsevier, pp. 413, p. 74-75.

Moschioni, G. (2002). A New Method for Measurement of Early Sound Reflections in Theaters and Halls. Instrumentation and Measurement Technology Conference, 2002. IMTC/2002. Proceedings of the 19th IEEE.

Müller, S. & Massarani, P. (2001). Transfer-Function Measurement with Sweeps. *Journal of the Audio Engineering Society,* **49**(6), p. 443-471.

Musha, T. & Kumazawa, T. (2008). Intensity Analysis Methods for Transient Signals. *Applied Acoustics,* **69** p. 60-67.

Navidi, W. (2006). *Statistics for Engineers and Scientists,* U.S.A. , McGraw-HIll, pp. 868, p. 157-170, 321-325, 772, 869.

Nelson, D. (1992). Impulse Response of very Small Rooms. AES 93rd Convention, San Francisco, CA. U.S.A. , October 1-4.

Newell, P. (2012). *Recording Studio Design, Third edition,* Oxford, U.K., Focal Press an Imprint of Elsevier, pp. 804, p. 133-186.

Noxon, A. M. (1992). Correlation Detection of Early Reflections. AES 11th International Conference: Test & Measurement.

Ohta, T., Yano, H., Yokoyama, S. & Tachibana, H. (2008). Sound Source Localization by 3-D Sound Intensity Measurement using a 6-channel microphone system Part 2: Application in room acoustics. *Internoise 2008 37th International Congress and Exposition on Noise Control Engineering,* p. 1-6.

Okubo, H., Otani, M., Ikezawa, R., Komiyama, S. & Nakabayashi, K. (2001). A System for Measuring the Directional Room Acoustical Parameters. *Applied Acoustics,* **62**(2), p. 203-215.

Olive, S. E. & Toole, F. E. (1989). The Detection of Reflections in Typical Rooms*. *Journal of the Audio Engineering Society,* **37**(7/8), p. 539-553.

Omoto, A. & Uchida, H. (2004). Evaluation Method of Artificial Acoustical Environment: Visualization of Sound Intensity. *Journal of Physiological Anthropology and Applied Human Science,* **23**(6), p. 249-253.

Oppenheim, A. V. (1989). *Discrete Signal Processing, Second edition,* New Jersey, U.S.A., Prentice Hall, pp. 870, p. 776-778.

Peltonen, T., Lokki, T., B, G., Merimaa, J. & Matti, K. (2001). A System for Multi-Channel and Binaural Room Response Measurements. AES 110th Convention, May 12–15, Amsterdam, The Netherlands., Jan 1.

Pihlajamäki, T. (2009). *Multi-resolution Short-time Fourier Transform Implementation of Directional Audio Coding.* M.Sc. Thesis, Helsinki University of Technology, Espoo, Finland, pp. 87.

Polack, J. (1996). Omni-directional Sound Source. WO Patent, patent application PCT/DK95/00399. patent number 96/011558. Sweden.

Protheroe, D., Ballagh, K., Prasad, D. & Conetta, R. (2013). Determination and Display of 3D Room Impulse Response. Reproduced Sound 2013, Nov. 12-14, Manchester, U.K. Proceedings of the Institute of Acoustics, **35** (2).

Pulkki, V. (2007). Spatial Sound Reproduction with Directional Audio Coding. *Journal of the Audio Engineering Society,* **55**(6), p. 503-516.

Raangs, R. (2005). *Exploring the Use of the Microflown.* Ph.D. Thesis, University of Twente, Enschede, The Netherlands, pp. 242.

Rechenberger, F. S. (2009). *Measurement System for the Localisation of Early Reflections.* M.Sc. Dissertation, University of Salford, Manchester, United Kingdom, Manchester, United Kingdom, pp. 1-187.

Reed, D. K. & Maher, R. C. (2009). An Investigation of Early Reflection's Effect on Front-Back Localization in Spatial Audio. AES 127th Convention 2009 October 9-12 New York NY, U.S.A. Audio Engineering Society,

Rettinger, M. (1968). The Statistics of Delayed Reflections*. *Journal of the Audio Engineering Society,* **16**(4), p. 436-439.

Roberts, T. W. (2009). *Design and Analysis of a Reflection Measurement System.* M.Sc. Dissertation, University of Salford, Manchester, United Kingdom, Manchester, United Kingdom, pp. 1-203.

Romero-Perez, J. (2010). Report of Angular Accuracy of the *Soundfield* ST350 for Detection of Direct Sound and First Reflection in Vertical Plane. *PhD Reports Salford,* p. 1-65.

Romero-Perez, J. (2011). Report of Measurements done with the p-p probe and the *Soundfield* Microphone, University of Salford, Manchester, U.K.

Romero-Pérez, J. (2011). Internal Evaluation Report titled: The use of Sound Intensity for Characterisation of Reflected Energy in Small Rooms, University of Salford, U.K.

Romero-Pérez, J. & Fazenda, B. M. (2009). Characterisation of the Directionality of Reflections in Small Room Acoustics. School of Computing and Engineering Researchers' Conference, University of Huddersfield, Dec 9th.

Romero-Perez, J., Fazenda, B. M. & Atmoko, H. (2009). Characterisation of Small Room Acoustics for Audio Production Reproduced Sound 25: The Audio Explosion, Nov 6, Brighton, U.K. Proceedings of the Institute of Acoustics, **31** (4).

Rubak, P. & Johansen, L. G. (2003). Coloration in Natural and Artificial Room Impulse Responses. AES 23rd International Conference, Copenhagen, Denmark, 2003 May 23-25.

Sage, R. J. (1975). *An Investigation of the Two-Microphone Method of Acoustic Intensity Measurement.* Bachelor Bs. Thesis, University of Southampton, Southamptom, United Kingdom, pp. 33.

Salomons, A. M. (1995). *Coloration and binaural decoloration of sound due to reflections.* PhD, Technische Universiteit Delft, The Nederlands, pp. 159.

Schlemmer, U. (2006). How do we perceive Early Reflexions? Some Notes on the Directivity of Music Instruments. 24Th Tonmeistertagung − Vdt International Convention, November, 2006, Oct 30.

Schroeder, M. R. (1987). Statistical Parameters of the Frequency Response Curves of Large Rooms*. *Journal of the Audio Engineering Society,* **35**(5), p. 299-306.

Schroeder, M. R. & Kuttruff, H. (1962). On Frequency Response Curves in Rooms. Comparison of Experimental, Theoretical, and Monte Carlo Results for the Average Frequency Spacing Between Maxima. *Journal of the Acoustical Society of America,* **34**(1), p. 76-80.

Sekiguchi, K., Kimura, S. & Hanyuu, T. (1992). Analysis of Sound Field on Spatial Information Using a Four-Channel Microphone System Based on Regular Tetrahedron Peak Point Method. *Applied Acoustics* **37**(4), p. 305-323.

Shin, K. & Hammond, J. K. (2008). *Fundamentals of Signal Processing: for Sound and Vibration Engineers,* West Sussex, U.K., John Wiley and Sons Ltd., pp. 416, p. 104, 333, 345.

Skålevik, M. (2011). Schroeder Frequency Revisited. Forum Acusticum, Aalborg, Denmark.

Slaney, M. (1993). An efficient implementation of the Patterson-Holdsworth auditory filter bank, Tech Report #35, Apple Computer Inc., U.S.A.

Smith, S. W. (1999). *The Scientist and Engineer′s Guide to Digital Signal Processing,* U.S.A., California Technical Publishing, pp. 664, p. 175, 365-367.

Spenceley, T. & Wiggins, B. (2009). Distance Coding and Performance of The Mark 5 And ST350 *Soundfield* Microphones and their Suitability for Ambisonic Reproduction. Proceedings of the Institute of Acoustics Reproduced Sound 25, Nov 19-20 Brighton, U.K. **39** (4).

Stan, G.-B., Embrechts, J.-J. & Archambeau, D. (2002). Comparison of Different Impulse Response Measurement Techniques. *Journal of the Audio Engineering Society,* **50**(4), p. 249-262.

Stanzial, D., Bonsi, D. & Prodi, N. (2000). Measurement of New Energetic Parameters for the Objective Characterization of an Opera House. *Journal of Sound and Vibration,* **232**(1), p. 193-211.

Stanzial, D., Bonsi, D. & Schiffrer, G. (2002). Four-dimensional treatment of linear acoustic fields and radiation pressure. *Acta Acustica united with Acustica* **88** p. 213-224.

Stanzial, D., Bonsi, D. & Schiffrer, G. (2005). Equipment for collection and measurement of quadraphonic sound data and metadata as well as a corresponding recording

procedure. EP Patent, patent application 04425909.1. patent number 1,538,876. France.

Stanzial, D. & Prodi, N. (1997). Measurements of Newly Defined Intensimetric Quantities and their Physical Interpretation. *The Journal of the Acoustical Society of America,* **102**(4), p. 2033-2039.

Stanzial, D., Prodi, N. & Schiffrer, G. (1996). Reactive Acoustic Intensity for General Fields and Energy Polarization. *Journal of the Acoustical Society of America,* **99**(4), p. 1868-1876.

Svend, G. (1985). Influence of Tripods and Microphone Clips on the Frequency Response of Microphones, Brüel and Kjaer, Volume 4, Denmark.

Tashev, J. I. (2009). *Sound Capture and Processing: Practical Approaches,* Wiltshire, U.K., John Wiley & Sons, pp. 365, p. 271.

Tervo, S. (2009). Direction Estimation Based on Sound Intensity Vectors. 17th European Signal Processing Conference (EUSIPCO).

Tervo, S. (2010). Estimation of reflective surfaces from continuous signals. *Acoustics Speech and Signal Processing IEEE,* p. 153-156.

Tervo, S. (2012). *Localization and Tracing of Early Acoustic Reflections.* Ph.D. Thesis, Aalto University, Espoo, Finland, pp. 143.

Tervo, S., Korhonen, T. & Lokki, T. (2011). Estimation of Reflections from Impulse Responses. *Building Acoustics,* **18**(1,2), p. 159-174.

Tervo, S., Patynen, J. & Lokki, T. (2009). Acoustic Reflection Path Tracing Using a Highly Directional Loudspeaker. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, WASPAA '09.

Thiele, R. (1953). Richtungsverteilung und Zeitfolge der Schallruckwurfe in Raumen. *Acustica,* **3** p. 291-302.

Thiele, R. & Meyer, E. (1956). Acoustical Investigations in numerous concert halls and broadcast studios using new measurement techniques *Acustica,* **6**(2), p. 425-444

Thrane, N., Frederlksen, E. & Thornsen, L. (1984). The Hilbert Transform, Brüel and Kjaer Technical Review, Volume 3, Denmark.

Thrane, N., Wismer, J., Konstantin-Hansen, H. & Gade, S. (1997). Practical use of the "Hilbert transform", Brüel and Kjaer Application notes, Denmark.

Tohyama, M. & Koike, T. (2008). *Acoustical Signal Processing,* U.K., Academic Press, pp. 334, p. 25, 30, 111, 127.

Toole, F. E. (2006). Loudspeakers and Rooms for Sound Reproduction—A Scientific Review*. *Journal of the Audio Engineering Society,* **54**(6), p. 451-476.

Toole, F. E. (2008). *Sound Reproduction: Loudspeakers and Rooms,* U.S.A., Focal Press an Imprint of Elsevier, pp. 569, p. 63, 69, 73, 75, 76, 78-80, 82-84, 110, 112, 128, 133, 138.

Torres-Guijarro, S., Pena, A., Rodríguez-Molares, A. & Degara-Quintela, N. (2011). A Study of Wideband Absorbers in a Non-Environment Control Room: Characterisation of the Sound Field by Means of p-p Probe Measurements. *Acta Acustica united with Acustica,* **97**(1), p. 82-92.

Trauth, M. H. (2010). Statistics on Directional Data. *MATLAB® Recipes for Earth Sciences.* Berlin Heidelberg, Germany: Springer-Verlag, p. 311-326.

Tuncer, E. & Friedlander, B. (2009). *Classic and Modern DOA Estimation,* Burlington Massachusetts, U.S.A., Academic Press and imprint of Elsevier, pp. 436, p. 189-190.

Van Honschoten, J. W. (2004). *Modelling and Optimisation of the Microflown.* Ph.D. Thesis, University of Twente, The Netherlands, pp. 224.

Van Lancker, E. (2000). Localization of Reflections in Auditoriums using Time Delay Estimation. AES 108th Convention, February 19-22 Paris, France.

Vaseghi, S. V. (2007). *Multimedia Signal Processing: Theory and Applications in Speech, Music and Communications*, John Wiley & Sons Ltd, pp. 1-677, p. 464-466.

Ville, J. (1948). Theorie et Applications de la Notion de Signal Analytique. *Cables et transmission,* **2**(1), p. 61-74.

Voetmann, J. (2007). 50 Years of Sound Control Room Design. *AES 122nd Convention Vienna Austria May 5-8 2007,* p.

Walker, R. (1995). Controlled Image Design: The Management of Stereophonic Image Quality, BBC Research and Development Report,

Walker, R. (1996). Room Modes and Low Frequency Responses in Small Enclosures. AES 100th Convention May - Copenhagen 1996 May 11-14.

Wallach, H., Newman, E. B. & Rosenzweig, M. R. (1973). The Precedence Effect in Sound Localization*. *Journal of the Audio Engineering Society,* **21**(10), p. 817-826.

Watkinson, P. & Fahy, F. J. (1984). Characteristic of Microphone Arrangements for Sound Intensity Measurement. *Journal of Sound and Vibration,* **94**(2), p. 299-306.

Wierzbicki, J., Małecki, P. & Wiciak, J. (2013). Localization of the Sound Source with the Use of the First-order Ambisconic Microphone. *Acta Physica Polonica Acoustic and Biomedical Engineering,* **123**(6), p. 1114-1117.

Yamasaki, Y. & Itow, T. (1989). Measurement of Spatial Information in Sound Fields by Closely Located Four Point Microphone Method. *Journal of the Acoustical Society of Japan.(E),* **10**(2), p. 101-110.

Yano, H., Ohta, T., Yokoyama, S. & Tachibana, H. (2008). Sound Source Localization by 3-D Sound Intensity Measurement using a 6-channel Microphone System Part 1 : Principle and Basic Applications. *Internoise 2008 37th International Congress and Exposition on Noise Control Engineering,* p. 1-6.

Yntema, D. R. (2008). *An Integrated Three-Dimensional Sound-Intensity-Probe.* Ph.D. Thesis, The University of Twente, Enschede, The Netherlands, pp. 180.

Yntema, D. R. & Van Honschoten, J. W. (2010). An Integrated 3D Sound Intensity Sensor using Four-Wire Particle Velocity Sensors: I. Design and Characterization. *Journal of Micromechanics and Microengineering,* **20** p. 1-7.

Yntema, D. R., Van Honschoten, J. W., De Bree, H. E. & Wiegerink, R. J. (2006). A Three Dimensional Microflown. *IEEE MEMS 2006, Istanbul, Turkey, 22-26 January 2006.,* p. 654-657.

Yost, W. (1974). Discriminations of interaural phase differences. *Journal of the Acoustical Society of America,* **55**(6), p. 1299-1303.

Index