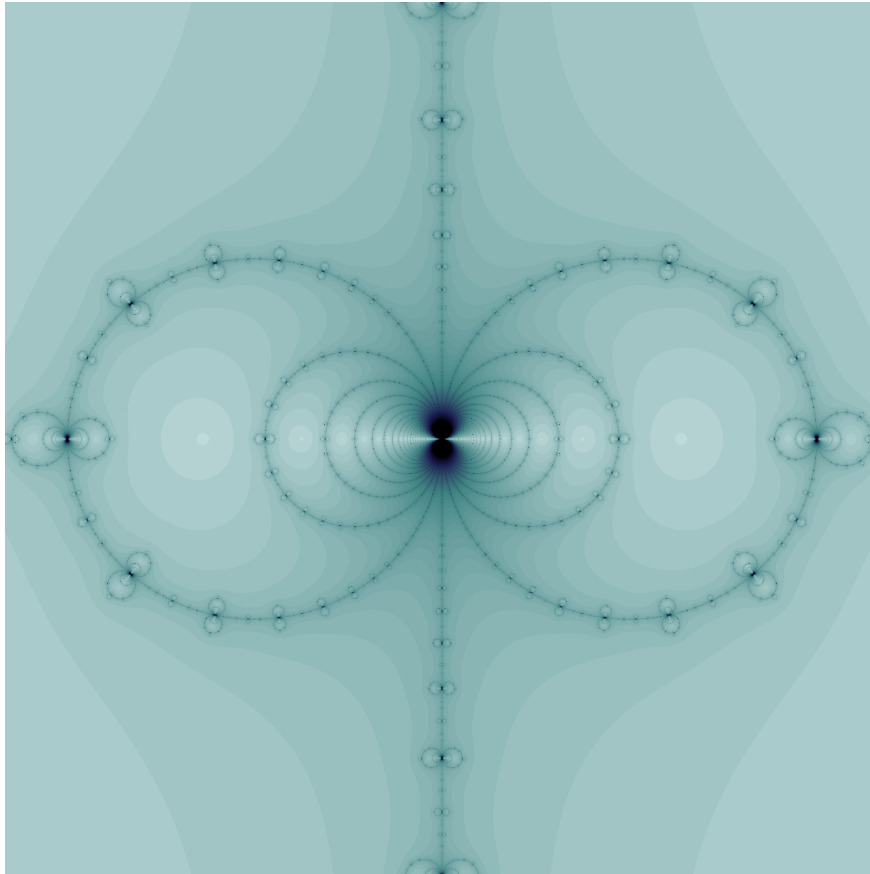




CHALMERS
UNIVERSITY OF TECHNOLOGY



Investigation of Local Variations of Room Acoustic Parameters

Master's thesis in Master Programme Sound and Vibration

GIULIO LAURENZI

MASTER'S THESIS ACEX30-18-107

Investigation of Local Variations of Room Acoustic Parameters

GIULIO LAURENZI



Department of Architecture and Civil Engineering
Division of Applied Acoustics
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2018

Investigation of Local Variations of Room Acoustic Parameters

© GIULIO LAURENZI, 2018.

Supervisor: Jens Ahrens, Department of Architecture and Civil Engineering

Examiner: Jens Ahrens, Department of Architecture and Civil Engineering

Master's Thesis ACEX30-18-107

Department of Architecture and Civil Engineering

Division of Applied Acoustics

Chalmers University of Technology

SE-412 96 Gothenburg

Telephone +46 31 772 1000

Cover: Newton Fractal created with MatLab script by Ankur Pawar

(<https://www.mathworks.com/matlabcentral/fileexchange/24840-newton-fractal>)

Typeset in L^AT_EX

Gothenburg, Sweden 2018

Investigation of Local Variations of Room Acoustic Parameters

GIULIO LAURENZI

Department of Architecture and Civil Engineering

Chalmers University of Technology

Abstract

In room acoustics the impulse response is an important piece of information because it is used in the determination of different acoustical parameters. It has been observed that the values of the parameters can change significantly, depending on the positions of the source and receiver transducer used in the measurements, even by a minimal position change.

The objective of this thesis is to investigate if these large variations around a given position are audible.

This project makes use of a continuous impulse response measurement system consisting of a microphone that is acquiring data through a constant movement in a defined trajectory. The data acquired is then processed to estimate the impulse responses with very high spatial resolution over the whole revolution of the microphone and consequently the room acoustic parameters are computed. The system has been tested in three different environments. The results are then compared with standard measurement data. These results show a good compatibility of the mean data.

An informal perceptual evaluation has been carried out to verify the perception of the parameter variations. While large parameter changes occurring between locations at a significant distance from each other are clearly audible, equally large parameter changes that occur locally are only partly audible. This suggests that some of the observed local variations are caused by inaccuracies of the measurement procedure.

Keywords: room acoustics measurements, continuous acquisition data in sound field synthesising, subjective perception, local variations of room acoustics parameters.

Acknowledgements

I sincerely thank my supervisor Jens Ahrens, thank you for all your patience, your time and the help you gave me to carry out this project.

I am very thankful for the support of Kilian Schulze-Forster who helped me with the processing of the sound-field microphone.

Thank you to all the personnel and students at the division of Applied Acoustics at Chalmers University of Technology for providing a really familiar and stimulating environment.

Finally thank you to my parents for their support and love.

Giulio Laurenzi, Gothenburg, September 2018

Contents

List of Figures	xi
List of Tables	xiii
1 Introduction	1
2 Theory	3
2.1 Analysis techniques	3
2.1.1 Impulse response	3
2.2 Room acoustics	3
2.2.1 Reverberation time	4
2.2.2 Early decay time (EDT)	4
2.2.3 Definition	5
2.2.4 Clarity	5
2.2.5 Centre time	5
2.2.6 Speech transmission index (STI)	6
2.2.7 Interaural cross correlation (IACC)	6
2.2.8 Lateral Energy Fraction (LEF)	7
2.3 Adaptive filter	7
2.3.1 Least Mean Square (LMS)	8
2.3.2 Normalised Least Mean Square (NLMS)	8
2.4 Impulse response modelling	9
2.5 Least Square Estimation	11
2.6 Human Auditory System	11
2.6.1 Just Noticeable Differences	11
2.6.2 Interaural Cues, Echo and Coloration	12
2.7 Sound-field Microphone	12
3 Methods	15
3.1 Measurement Simulation	15
3.2 Estimation of the Impulse Responses	17
3.3 Extension of the Room Impulse Response Beyond its Noise Floor	17
3.4 Computation of the Objective Parameters	19
3.4.1 Interaural Cross Correlation Assumptions	19
3.4.2 Reverberation Time Computation	22
3.5 Measurements	24
3.5.1 Moving Microphone Measurements	24

3.5.2	Static Measurements - Directional Source	29
3.6	Perceptual Evaluation	29
4	Results	33
4.1	Averaged Results - Omnidirectional Sources	33
4.2	Local Parameters - Directional Source	36
4.3	Data Comparison From Moving Microphone and Static Microphone Measurements on Equal Positions	50
4.4	Perceptual Evaluation	54
5	Conclusion and Discussion	55
	Bibliography	57

List of Figures

2.1	Image source	9
3.1	Partitioned convolution [17]	16
3.2	Comparison: measured signal and error signal, respective IR and backward integration	18
3.3	Filter Bank Transfer Function Magnitudes	19
3.4	Result of the process of the extend IR	20
3.5	Magnitude squared coherence of different set up of a pair of micro- phones at 17 cm distance from each other	21
3.6	Magnitude squared coherence for different microphones according to Brandstein [21]	22
3.7	Correlation for two omnidirectional microphones and HRIR with band limited noise	23
3.8	B-Format microphone - Core Sound TetraMic	25
3.9	VariSphear - [24]	25
3.10	Comparison of the filtered and unfiltered B-format microphone [25]. .	27
3.11	Position of the VariSphear and loudspeaker in the different Labs. . . .	28
4.1	Audio lab parameters comparison	34
4.2	Reverberant chamber parameters comparison	35
4.3	Comparison of the result with different resolution over the measure- ment trajectory - B format mic.	36
4.4	Early decay time - EDT	39
4.5	Reverberation Time T_{30}	40
4.6	Comparison of the Schroeder's curve in the Listening room at 1 kHz .	41
4.7	Comparison of the Schroeder's curve	41
4.8	Clarity C_{50}	43
4.9	Definition D_{50}	44
4.10	Centre Time T_s	46
4.11	Speech transmission index STI	47
4.12	Interaural cross correlation - IACC	49
4.13	Lateral energy fraction	50
4.14	Audio Lab centre - static VS moving microphone comparison	52
4.15	Reverberation Chamber - static VS moving microphone comparison .	53

List of Tables

2.1	Just Noticeable Difference according to ISO 3382-1:2006	12
3.1	Measurements overview	24
3.2	Positions of the IR used in the perceptual evaluation test.	31

1

Introduction

The room acoustics' parameters, derived from the computations based on the energy relations of the sound, give important data to the evaluation of the acoustic behaviour in a specific environment. The room impulse response (RIR), which represents the synthesis of the sound field, is the main information used to compute the parameters. In [1], it is reported that the room acoustics measurements show variations depending on the geometry of the room and, the position of the source and of the microphone. These differences could be significant even with small relocations of the source or the receiver. This is proved in the studies of Nielsen *et al.* in [2], where the variations of clarity C_{80} measured on a single sitting place on a concert hall, present data fluctuations within the limit of the just noticeable detection. According to this behaviour, in room acoustics, to get a realistic representation of the considered space, it is suggested to take measurements over a large number of positions. This technique would return an averaged evaluation but it does not take into consideration the capacity of the human auditory system to suppress noise, reverberance and sound coloration [4]. This means that, in different cases, the parameters' data cannot determine what is really perceived.

De Vries *et al.* in [3] investigated the variations of the apparent source width by studying the perception of the lateral energy fraction and interaural cross correlation derived from measurements and simulations with a microphone array, where every transducer is placed at a distance of 0.05 m to each other, over the full width of the room under study. The measurements of this study return relevant fluctuations of the parameters even at close position. In addition, it is revealed that the measured fluctuations were not clearly perceived as differences. According to this, the author asserts that more studies will be required to be able to have a clear comprehension of the perception results.

The aim of this master thesis is to investigate the variations of the room acoustic parameters, across distant and close positions, with focus on the later, to verify whether they are audible. The parameters taken into consideration are not only the ones analysed in [2] and [3], but also definition, speech transmission index, centre time, early decay time and reverberation time. To do this, a measurement system consisting of a moving microphone following a circular trajectory with constant speed has been implemented. The acquisition of the data is not done position after position, it happens continuously, with a uniform movement which does not stop during the measurement. The procedure consists of synthesising the sound field with very high spatial resolution over a circular trajectory of the microphone, computing the room acoustic parameters and verifying the reliability of the results while com-

paring them with standard measurement results and finally studying the perception of the sound field at crucial positions related to the computed data.

Techniques of sound field measurement with moving microphones have already been studied and implemented, as the technique described by Thibaut Ajdler in [5]. Another technique was presented by Fabrice Katzberg in [6]. The technique used in this thesis is designed to perform head related impulse response measurement, introduced by Gerard Enzner in [7]. It has been chosen for its simplicity and ease of applicability with the devices present at the university division. This technique will process the acquired data over a circular revolution, estimating the room impulse response over the trajectory with an azimuth angle interval which can be set to have large variation in degrees, up to a decimal of a degree.

The outline of this report is structured as follows: the second chapter contains the mathematical and theoretical background of the different processes used in the project which may be needed in order to understand the methods showed in this report. The Method is presented in chapter three, in which the different steps adopted in the project will be explained, from the simulation to the final results. Chapter four shows the results of the different measurements and finally the Conclusion section, which ends the thesis report, summarises the important discoveries.

2

Theory

In order to understand the following parts of this project, the main principles of room acoustics and mathematical background have to be recalled. These basics include the definition of the most important acoustic parameters; the theory about the adaptive filters; the *image-source model* to generate virtual impulse responses and some relevant information about the human hearing.

2.1 Analysis techniques

2.1.1 Impulse response

The impulse response (IR) $h(t)$ contains all the information about how a linear time-invariant (LTI) system reacts to an input signal (an impulse). The output of the system $y(t)$ related to the input $x(t)$ can be determined by the convolution process, in which the input is convolved with the IR:

$$y(t) = \int_{-\infty}^{\infty} x(\tau) \cdot h(t - \tau) d\tau = x(t) * h(t). \quad (2.1)$$

A room can be considered as an LTI system even if in reality it can have time-varying acoustic properties like air movement or temperature changes. Usually the changes of these conditions are so slow that it can be considered stable during the measurement process [1].

Starting from the IR, which in room acoustics represents the synthesis of the sound field, many important pieces of information can be derived regarding sound propagation in a room under test. The impulse response is also a function of the position of the source and receiver and can varies greatly with their position inside the room. Once the impulse response of a room has been measured, the process of playing a sound in that room can be simulated by convolving the sound with the room's impulse response. The output of this operation is very similar to the measured response in the real case, with the same conditions of the source and receiver.

2.2 Room acoustics

In a closed space, like a room, the sound waves from a source are subjected to reflections. The geometries of the room, the material of the walls and the objects present in the room are all dependent factors which may not cause an immediate decay of the sound energy. All the reflections, summed to the direct signal from the

source to the receiver will contribute to the resulting perceived sound. In the case of the energy in the room decaying slowly (long reverberations in the room), it could become hard to distinguish and understand speech.

2.2.1 Reverberation time

The reverberation time is one of the most important and distinguishable acoustic parameters, as it gives a general idea of the acoustic behaviour of the room. It defines the time, in seconds, it takes for energy of the sound field to drop by 60 dB from its steady-state level [8].

The traditional method of measurement consists of the evaluation of the decay curve. To obtain the decay curve, a room is excited by random noise to steady-state conditions, until at a certain moment in which the excitation is interrupted and a recording device starts to record the decay process. This process contains fluctuations due to the type of noise used, so it is required to perform more measurements to average them all into one decay curve. Because the noise floor can result in being too high to correctly record a fall of 60 dB in the sound field, instead the T_{30} , the time needed for the sound field to decay by 30 dB, is used to calculate the fall of 60 dB, by extending the slope of its curve; T_{20} , instead, consider the time needed for the sound field to decay by 20 dB.

A more accurate method to measure the reverberation time is the *backward integration*, introduced by Schroeder in 1965. It is computed from the impulse response of the room and it consists of the relation between the impulse response $h(t)$ and the average squared of all the possible decay curves $\langle g^2(t) \rangle$ [8] as shown in Eq. 2.2

$$\langle g^2(t) \rangle = \int_t^\infty [h(x)]^2 dx = \int_0^\infty [h(x)]^2 dx - \int_0^t [h(x)]^2 dx \quad (2.2)$$

The slope of the straight line is the decay rate d , in decibels per second, then the reverberation time is computed as:

$$RT = 60/d \quad (2.3)$$

Eq. 2.3 can be used for calculating T_{30} and T_{20} . Furthermore the standard ISO 3382 defines the evaluation range from 5 dB to 35 dB below the steady-state level for T_{30} . For T_{20} the evaluation range is from 5 dB to 25 dB. In all cases, for a maximum underestimation of 5 percent, the level of the background noise must be at least 15 dB below the evaluation range of the impulse response.

2.2.2 Early decay time (EDT)

The direct sound of a source, in a room, is followed by a dense series of first order reflections. In this range of time the acoustical defects of the room appear. According to [8], the early decay time defines the time interval in which decay from 0 to -10 dB is obtained, then multiplied by a factor of six.

2.2.3 Definition

As presented by H. Kuttruff in [8] the *definition* is a parameter to define an objective criterion to measure the distinctiveness of the sound using the IR. Both integrals of Eq. 2.4 include the direct sound. A definition with value of 1 means that the IR does not contain energy due to reflection in the time period after the first 50 ms. D_{50} is a useful description of speech intelligibility.

$$D_{50} = \frac{\int_0^{50ms} [h(t)]^2 dt}{\int_0^{\infty} [h(t)]^2 dt} \cdot 100\% \quad (2.4)$$

2.2.4 Clarity

The Clarity is a quantity which is similar to D_{50} but it will characterise the transparency of music in concert halls [8]. The difference with the definition is the higher limit of delayed energy (80 ms) and the comparison with the rest of the energy (from 80 ms to ∞) instead of the whole energy. This variation is due to the assumption that with music signals the reflections are less detectable than in speech signals.

$$C_{80} = 10 \log_{10} \frac{\int_0^{80ms} [h(t)]^2 dt}{\int_{80}^{\infty} [h(t)]^2 dt} dB \quad (2.5)$$

The result of clarity is expressed in decibel, a value of 0 dB will be considered sufficient for fast musical passages, a value of -3 dB can be still considered tolerable.

The equation for C_{50} is the same as the C_{80} with a different limit of early delayed energy: 50 ms instead 80 ms.

2.2.5 Centre time

The use of a defined constant time limit (50 ms and 80 ms) separating the early and late energy in D_{50} and C_{80} is an approximation of how the human hearing system processes the reflections. In some critical cases, a small change in the delay time of a strong reflection can result in a considerable variation in the value of the Clarity and Definition. A remedy to this is with the Centre time T_s in Eq. 2.6 which does not contain a sharp time limit [8]. It describes the centre of gravity of the squared impulse response. Low values of T_s will correspond to a high speech intelligibility.

$$T_s = 10 \log_{10} \frac{\int_0^{\infty} t \cdot [h(t)]^2 dt}{\int_0^{\infty} [h(t)]^2 dt} ms \quad (2.6)$$

2.2.6 Speech transmission index (STI)

The speech intelligibility is of focus in the environments created for lyrical music performances and speech. There are different parameters to measure it, like the Clarity (C_{80}) and the Definition (D_{50}). Currently the most reliable measurement for speech intelligibility is the speech transmission index.

According to the European standards [9] the STI can be calculated with two methods.

The first one, the *direct method* is based on the assumption that the speech can be considered as an amplitude modulated signal in which the degree of modulation carries the speech information. If the noise and reverberation are added in the transmission path, the degree of modulation will be reduced, with the consequence of a reduced intelligibility. The modulation transfer function (MTF) is measured through the emission of noise in the octave bands from 125 Hz to 8 kHz. Each of these octave bands are modulated with 14 different modulation frequencies. The modulation reduction factor is obtained by computing the ratio between the original and the received degree of modulation for each of the 98 possible combinations. Finally, the modulation reduction factor is weighted and averaged to obtain a value between 0 and 1, where 0 corresponds to very poor and 1 to excellent STI.

The second method, the *indirect method*, makes use of the room IR to calculate the modulation reduction factor. To do this the following equation is used:

$$m_k(f_m) = \frac{\left| \int_0^\infty h_k(t)^2 e^{-j2\pi f_m t} dt \right|}{\int_0^\infty h_k(t)^2 dt} \cdot \frac{1}{1 + 10^{-SNR_k/10}} \quad (2.7)$$

where $m_k(f_m)$ is the MTF, $h_k(t)$ the impulse response of octave band k , f_m the modulation frequency, SNR_k is the signal to noise ratio in dB. Note that this equation in IEC 60268-16:2011 has a typographical error: it omits the squaring of the impulse response in the numerator of the first term.

With the determined modulation transfer function the effective SNR is then computed which is, in turn, used to calculate the transmission index (TI). The derived TI_{k,f_m} are then averaged over the modulation frequencies to obtain the modulation transfer index MTI_k per octave band.

Finally the MTI_k are weighted with the *weight factor* and the *redundancy factor* and summed to return the STI.

2.2.7 Interaural cross correlation (IACC)

According to [8], *spaciousness* or spatial impression is caused by reflections which reach the listener from lateral directions. Assuming that a sound wave is reaching the listener from the front producing an equal sound pressure at both ears, whereas the same wave is reaching the listener from another angle producing different signals at the ears, it is possible to define the laterally reflected sound from these differences. To measure this, the cross-correlation function is used. This measures the similar-

ity between two signals at the two receiving points (P_1 and P_2). The correlation coefficient R may vary between $+1$ and -1 , when R is 0, then the two signals are completely uncorrelated.

$$R = \frac{\overline{P_1(t) \cdot P_2(t)}}{\sqrt{\overline{P_1^2} \cdot \overline{P_2^2}}} \quad (2.8)$$

Adapting Eq. 2.8 to the human head where the IRs h_L and h_R are measured at the left and right ear, results in Eq. 2.9. Since these signals are transient, time averaging is not necessary, but they will be integrated over time:

$$R(\tau) = \frac{\int_{t1}^{t2} h_R(t) h_L(t + \tau) dt}{\sqrt{\int_{t1}^{t2} [h_R(t)]^2 dt \int_{t1}^{t2} [h_L(t)]^2 dt}} \quad (2.9)$$

Where $t1$ is 0 and $t2$ is 100 ms to limit the integration to the early reflections. The interaural cross correlation (IACC) is obtained by the maximum of the correlation coefficient R in the range of $|\tau| < 1$ ms.

2.2.8 Lateral Energy Fraction (LEF)

Another parameter to describe the *spaciousness* suggested in [8] is the Lateral Energy Fraction. This measurement considers the contribution of the reflections in the range between 5 and 80 ms and the relative incoming angle, assuming that the listeners head is facing the source. The data of LEF are averaged over the four octave bands of 125 Hz, 250 Hz, 500 Hz and 1000 Hz because the low- and mid-frequency components contribute most to the sensation of spaciousness.

$$LEF = 100 \left(\frac{\int_{5ms}^{80ms} [h_L(t)]^2 dt}{\int_0^{80ms} [h_T(t)]^2 dt} \right) \quad (2.10)$$

where $h_L(t)$ corresponds to the impulse response measured with a bidirectional microphone and $h_T(t)$ is the impulse response measured using an omnidirectional microphone.

2.3 Adaptive filter

A filter is a tool used to extract or enhance desired information contained in a signal. An implementation of this is the adaptive filter which uses an adaptive algorithm to update the filter coefficients so that the filter can be used in an unknown and changing environment. The adaptive algorithm determines the filter characteristics by adjusting the filter coefficients according to the signal conditions and performance criteria. An example of performance criterion is based on an error signal, which is the difference between the output signal of the filter and a given reference signal.

2.3.1 Least Mean Square (LMS)

An efficient way of calculating the optimal filter is by using the LMS algorithm, as presented in [10]. Using this algorithm, the filter goes through a recursive adaptation for every new input sample, $x(n)$, and its corresponding desired output sample, $d(n)$, so that the error signal $e(n)$ is minimised in the mean-square sense.

The three steps to complete each iteration of the LMS algorithm are presented as follows:

In Eq. 2.11 the calculation of the filter output $y(n)$ is presented, where $\mathbf{h}(n)$ is the impulse response vector of the filter and $x(n)$ the input sample.

$$y(n) = \mathbf{h}^T(n)\mathbf{x}(n). \quad (2.11)$$

The estimated error $e(n)$ is calculated according to Eq. 2.12

$$e(n) = d(n) - y(n) \quad (2.12)$$

and finally in Eq. 2.13 the adaptation recursion is represented, where μ is a parameter that permits to minimise the estimated error.

$$\mathbf{h}(n+1) = \mathbf{h}(n) + 2\mu \cdot e(n)\mathbf{x}(n). \quad (2.13)$$

2.3.2 Normalised Least Mean Square (NLMS)

NLMS algorithm is a special implementation of the LMS algorithm that takes into account the variation of the signal level at the filter input and selects a normalised step-size parameter μ , which results in a stable and fast converging adaptation algorithm. Starting from Eq. 2.13, $\mu(n)$ is selected so that the a posteriori error is minimised in magnitude:

$$e^+(n) = d(n) - \mathbf{h}^T(n+1)\mathbf{x}(n), \quad (2.14)$$

Substituting Eq. 2.13 in Eq. 2.14 and rearranging, it is obtained:

$$e^+(n) = [1 - 2\alpha\mathbf{x}^T(n)\mathbf{x}(n)]e(n) \quad (2.15)$$

where α is the convergence coefficient used to scale the momentary error.

Minimising $(e^+(n))^2$ with respect to $\mu(n)$ results in the following equation:

$$\mu(n) = \frac{1}{2\mathbf{x}^T(n)\mathbf{x}(n)}, \quad (2.16)$$

which forces $e^+(n)$ to zero. Substituting equation 2.16 in equation 2.13, the NLMS recursion is obtained:

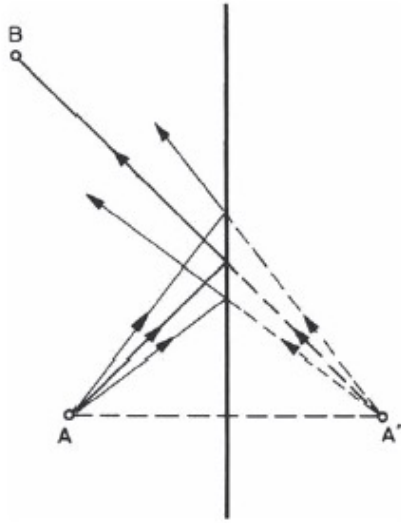
$$\mathbf{h}(n+1) = \mathbf{h}(n) + \frac{1}{2\mathbf{x}^T(n)\mathbf{x}(n)}e(n)\mathbf{x}(n). \quad (2.17)$$

2.4 Impulse response modelling

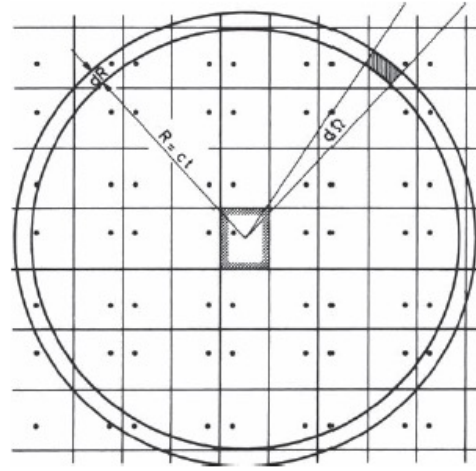
The sound propagation is described by the wave equation, with which it is possible to obtain the impulse response of a room. The image source model is one of the existing techniques for the computational modelling of room impulse response.

The image source method described in this section will only explain the simple case of rectangular, *box-shaped* rooms, as it is the same case used in the project. As presented in [8] and [11], this technique consists in representing the reflection of the source A in front of a rigid wall, to the receiver B . The virtual source A' is symmetrically located behind the wall, on the line perpendicular to the wall, parallel to the original source A and at the same distance from the wall as the original source A (see Figure 2.1a). It is assumed that the image source emits the same sound signal as the original source, but the energy can be reduced depending on the absorption coefficient α of the wall.

The concept of image source could be considered exact only in the case when the impedance of the reflecting boundary is $+1$ or -1 , because the absorption coefficient is defined only for plane waves.



(a) Construction with one wall [8]



(b) case of rectangular room (2D) [8]

Figure 2.1: Image source

Considering a single frequency point source at position A with coordinates (x_A, y_A, z_A) and receiver B with coordinates (x_B, y_B, z_B) , $R_1 = (x_A - x_B, y_A - y_B, z_A - z_B)$ is the distance of the two points. The pressure in free field will correspond to:

$$p(\omega, A, B) = \frac{e^{(j\omega(\|R_1\|/c - t))}}{4\pi R_1}, \quad (2.18)$$

Applying the Fourier transform, the IR is obtained:

$$h(t) = \frac{\delta\left(t - \frac{\|R_1\|}{c}\right)}{4\pi\|R_1\|} \quad (2.19)$$

Placing a wall in front of the source and adding the image source, the pressure at the receiver will result from the sum of the two sources:

$$h(t) = \frac{\delta\left(t - \frac{R_1}{c}\right)}{4\pi R_1} + \frac{\delta\left(t - \frac{R_2}{c}\right)}{4\pi R_2} \quad (2.20)$$

Where R_2 is the distance between B and A' .

Considering the three perpendicular walls in the x , y and z directions, seven additional virtual sources will be considered giving the following equation for the IR:

$$h(t) = \sum_{i=1}^8 \frac{\delta\left(t - \frac{\|R_i\|}{c}\right)}{4\pi\|R_i\|} \quad (2.21)$$

Where $R_i = (x_{A_i} \pm x_B, y_{A_i} \pm y_B, z_{A_i} \pm z_B)$ is the distance between the source A_i and receiver B .

Then assuming that each image is itself imaged, because the sound between two parallel rigid walls is infinitely reverberated (see figure 2.1b), then each of the first eight sources will be repeated with a period of $2L_x$, $2L_y$ and $2L_z$ where L_x , L_y , L_z are the dimensions of the room. This will result in:

$$h(t) = \sum_{i=1}^8 \sum_{r=-\infty}^{\infty} \frac{\delta\left(t - \frac{\|R_i + R_r\|}{c}\right)}{4\pi\|R_i + R_r\|}, \quad (2.22)$$

Where r is the integer vector triplet (n, l, m) and $R_r = 2(nL_x, lL_y, mL_z)$.

In the case that the walls are not totally reflective, it will be assumed that the acoustic property of each surface by its reflection coefficient β is itself related to the absorption coefficient α according to: $\alpha = 1 - \beta^2$. The reflection coefficients for each surface will be denoted with $\beta_{x,d}$, $\beta_{y,d}$ and $\beta_{z,d}$ where $d = 1, 2$, the sub index 1 corresponds to wall closest to the origin and sub index 2 corresponds to the opposing wall. The IR function with the introduction of the wall absorption will become:

$$h(t) = \sum_{p=0}^1 \sum_{r=-\infty}^{\infty} G_{p,r} \frac{\delta\left(t - \frac{\|R_p + R_r\|}{c}\right)}{4\pi\|R_p + R_r\|}, \quad (2.23)$$

where $G_{p,r}$ is the amplitude factor, R_p is the distance between the source and receiver, now expressed in the terms of the vector $p = (u, v, w)$ as $R_p = (x_A - x_B + 2ux_B, y_A - y_B + 2vy_B, z_A - z_B + 2wz_B)$.

The attenuation factor is defined as follows:

$$G_{p,r} = \beta_{x,1}^{n-u} \beta_{x,2}^{n-l} \beta_{y,1}^{l-v} \beta_{y,2}^{l-l} \beta_{z,1}^{m-w} \beta_{z,2}^{m-l} \quad (2.24)$$

2.5 Least Square Estimation

In presence of different observations, it is interesting to represent the data with something more meaningful as a function which fits the data and gives the trend of the observations. The series of observations can be represented by the sum of m functions with n coefficients C , which defines as:

$$\sum_{j=1}^n X_{i,j} C_j = y_i, \quad (2.25)$$

where $i = 1, 2, \dots, m$, y is the function and X the observations.

One of the most popular methods used for this purpose is the least square method [12], it consists of finding the minimal squared vertical distance (error) between the observed data and the approximated curve, finding the minimum value of the square of Eq. 2.25:

$$\sum_{j=1}^n |X_{i,j} C_j - y_i|^2. \quad (2.26)$$

To solve the least square it is necessary to calculate the derivative of the function in respect of every coefficient and create a system of equations. An equivalent method to solve the least square, is writing 2.26 in the matrix form $\mathbf{XC} = \mathbf{y}$ where,

$$\mathbf{X} = \begin{bmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,n} \\ X_{2,1} & X_{2,2} & \dots & X_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m,1} & X_{m,2} & \dots & X_{m,n} \end{bmatrix}, \mathbf{C} = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

and solve the normal equation:

$$\mathbf{X} \mathbf{X}^T \hat{\mathbf{C}} = \mathbf{X}^T \mathbf{y} \rightarrow \hat{\mathbf{C}} = \frac{\mathbf{X}^T \mathbf{y}}{\mathbf{X} \mathbf{X}^T} \quad (2.27)$$

2.6 Human Auditory System

The human auditory system is a complex set of organs. This section will outline a brief introduction to some phenomena relevant to this project.

2.6.1 Just Noticeable Differences

Studies have been performed to determine by how much the value of a parameter needs to change so the difference can be perceived by the human hearing system. These data is called Just Noticeable Differences (JND).

Some of these JND data are present in the standard ISO 3382-1 as shown in Table

2.1, but some research propose even other data. In the case of Clarity and Speech Transmission Index, Bradley, Reich and Norcross [13] state that people with trained critical listening can recognise a difference of 0.9 dB in C_{80} , 1.1 dB in C_{50} and 0.03 in STI. For a case of a non-trained ear, the JND will result in 3 dB for C_{50} and 0.1 for STI.

Regarding to the reverberation time, M. Blevins in [14], defines the JND to be about the 24% of the referenced RT.

Table 2.1: Just Noticeable Difference according to ISO 3382-1:2006

Acoustic quantity	Single number frequency averaging (Hz)	JND
EDT in s	500 to 1000	Rel. 5%
C_{80} in dB	500 to 1000	1 dB
D_{50}	500 to 1000	0.05
Ts in ms	500 to 1000	10 ms
$IACC$	/	0.075
LF	125 to 1000	0.05

2.6.2 Interaural Cues, Echo and Coloration

To understand how the human auditory system is able to localise sound sources in space, it must be considered that the arrival times of the sound wave emitted from a single source are not exactly the same at the left and right ear due to the different path-lengths to both ears [4]. This difference is called *interaural time difference* (ITD). The presence of the head between the ears does not only determine the difference in time travel to the ears, but also causes attenuation of the sound wave at the ear opposite to the source, which leads to *interaural level difference* (ILD) of both ear signals.

In a complex listening situation, where room reflections are present, the sound from different directions can simultaneously reach the listener. This result of the superposition of all the reflections is translate into a coloration of the perceived sound due to the comb filter effect. Depending on the level of the reflection and its time delay, the reflection can be perceived as an echo [8].

2.7 Sound-field Microphone

A tetrahedral array of cardioid microphone capsules is a type of sound-field microphone. With this type of microphone it is possible to perform Ambisonic sound recordings which represent a three-dimensional sphere of the sound field. The four unprocessed signals recorded from the tetrahedral microphone, called A-format signal, are then converted to a standard Ambisonic/B-format signal. The B-format signal consists of four channels which represents the omnidirectional information (W) and the figure of eight information for front/back (X), left/right (Y) and up/down (Z) directions [15]. Defining the four capsules of the tetrahedral microphone and relative A-format signals as:

- Left front up (LFU);
- Right front down (RFD);
- Left back down (LBD);
- Right back up (RBU);

The four B-format signals are generated according Eq. 2.28:

$$\begin{aligned}W &= LFU + RFD + LBD + RBU, \\X &= LFU + RFD - LBD - RBU, \\Y &= LFU - RFD + LBD - RBU, \\Z &= LFU - RFD - LBD + RBU.\end{aligned}\tag{2.28}$$

In this project the sound field microphone will be called B-format microphone, to avoid confusion with the omnidirectional component of the sound-field microphone and usual omnidirectional microphone.

3

Methods

This section presents the methodology applied to simulate and perform the measurements and describes the post-processing procedures applied to the measured data to obtain the room acoustics parameters.

3.1 Measurement Simulation

Before taking measurements, the whole process has been tested within a simulated environment to validate the implementation. In order to simulate a measurement of a moving microphone, the room impulse response (RIR) at any point in the revolution of the microphone has to be known. The image-source method presented in 2.4 is used for the simulation of the RIRs. This method is based on the creation of virtual sources in order to simulate the reflections of the walls of a room and it is simple to implement for a box-shaped room.

Initially a slightly modified version of the Allen and Berkly [11] implementation has been used. The use of integer propagation time for the calculation of the reflections, in this particular implementation, generated errors in the simulation. This will not respect the causality necessary for the IR estimation process (see further 3.2). To avoid the problem, Lehmann and Johansson's implementation [16], which uses a fractional propagation time has been adopted.

The positions of the RIRs are calculated for each sample at the azimuthal interval $\Delta\theta$ obtained by dividing the circular trajectory of the microphone in many points, by the revolution time over the complete circle T_{360} multiplied with the sampling frequency f_s . This interval assumes a constant revolution velocity as required by the estimation process.

$$\Delta\theta = \frac{2\pi}{T_{360} \cdot f_s} \quad (3.1)$$

Successively the computed RIRs will be convolved with the excitation signal which consists of a Maximum Length Sequence (MLS). This convolution is performed sample by sample, iterating on the samples of the MLS signal. To manage the convolution of the excitation signal with the different IRs introduced at each position/sample, the partition convolution method described by Armelloni, Giottoli and Farina in [17] based on a *overlap-and-save* standard process has been used. It consists of the division of every IR in equally sized blocks of K samples. In this case the length of K corresponds to one sample. Each one of these segments will be considered as separate impulse responses defined as S_P , where P is the quantity of the

blocks. Every block is zero-padded to the length L which corresponds to the power of 2 in which its value is closer to $2K$ (in this implementation $L = 2$), and processed with the Fast Fourier Transform (FFT). The excitation signal has been divided in blocks of length L , in which the FFT is applied and then multiplied with every segment S_P , creating different segments S for every block. The resulting segments are then summed and, at the end, reconverted to the time domain through the inverse FFT process. Only the last $L - K$ samples of every block are kept (only the last sample in this implementation). The process is presented with a block diagram in Figure 3.1.

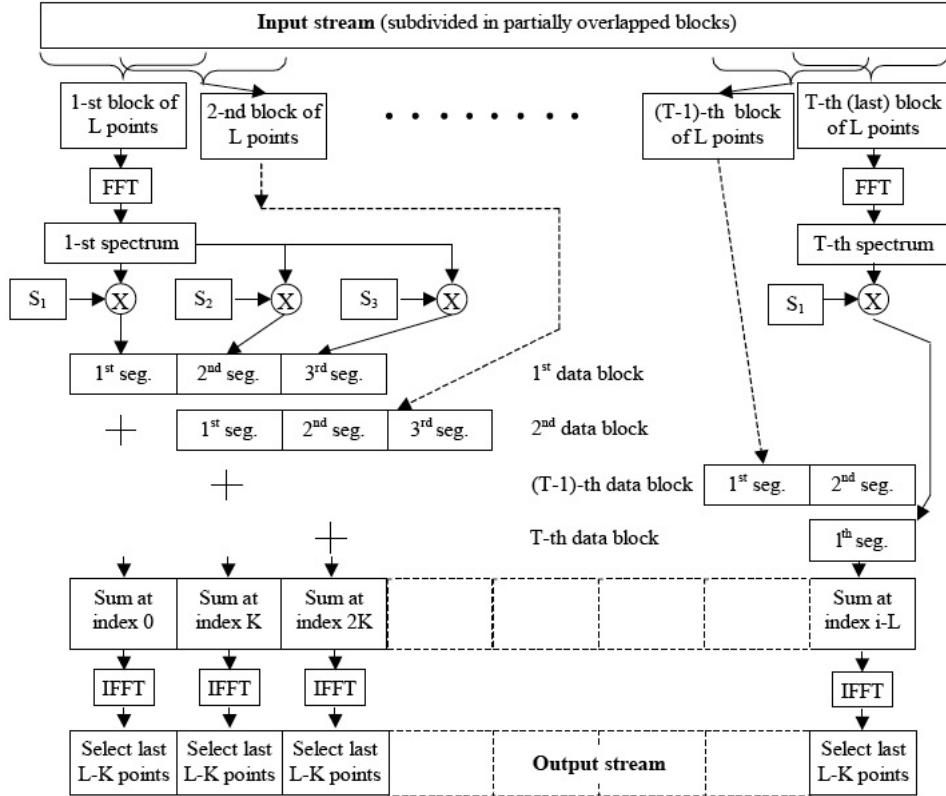


Figure 3.1: Partitioned convolution [17]

The excitation signal used, the MLS, will always be generated with the same sequence of samples, giving that the output is known. This will simplify the test process and will be helpful to the satisfaction of the causality required for the estimation process. The sampling frequency used in the simulation is 24 kHz which permits to have measurements up to the 8 kHz octave band. The radius of the microphone revolution used in the simulation corresponds to 1 meter. To obtain relevant results with the estimator, different revolution times of the microphone have been tested.

3.2 Estimation of the Impulse Responses

The data acquired needs to be processed to estimate the impulse responses at different positions. This process has been performed using the technique presented in [7], created for the measurement of head related impulse response (HRIR).

The HRIR measurement, done with Enzner's method, is performed with a moving microphone, through a circular trajectory, where the revolution of the microphone must be at a constant velocity.

The result of the recorded data is given by the convolution model:

$$y(t) = \sum_{\tau=0}^N h(\tau, \theta_t) x(t - \tau) + n(t) \quad (3.2)$$

where $h(\tau, \theta_t)$ is the time-varying IR at azimuth $\theta_t = \omega t \frac{1}{f_s}$, f_s is the sampling frequency, ω corresponds to $\frac{2\pi}{T_{360}}$, T_{360} is the revolution time. N denotes the length of the IR and $n(t)$ the noise. In this way every angle θ_t corresponds to a time index t .

The identification of every IR is based on the NLMS adaptive filtering algorithm. Generating the estimation of the impulse response $\hat{h}$ as:

$$\hat{h}(\theta_{t+1}) = \hat{h}(\theta_t) + \mu_0 \frac{e(t)x(t)}{\|x(t)\|^2} \quad (3.3)$$

where μ is the step-size parameter to minimise the estimation error $e(t)$ which correspond to:

$$e(t) = y(t) - \hat{h}^T(\theta_t)x(t) \quad (3.4)$$

To obtain good results from this technique, a causal relationship is required between the reference signal and the recorded signal. This means that it is necessary to know the initial delay in samples between the reference and the measured signal. Another information required is the total length of the final estimated IR.

3.3 Extension of the Room Impulse Response Beyond its Noise Floor

Despite the estimation function being supplied with demo files, which produces results with an estimation error signal approximately lower compared to the input signal (see Figure 3.2a), the impulse responses resulting from the process present a high noise floor. This will contribute in giving erroneous results in the computation of the room acoustic parameters with the impulse responses.

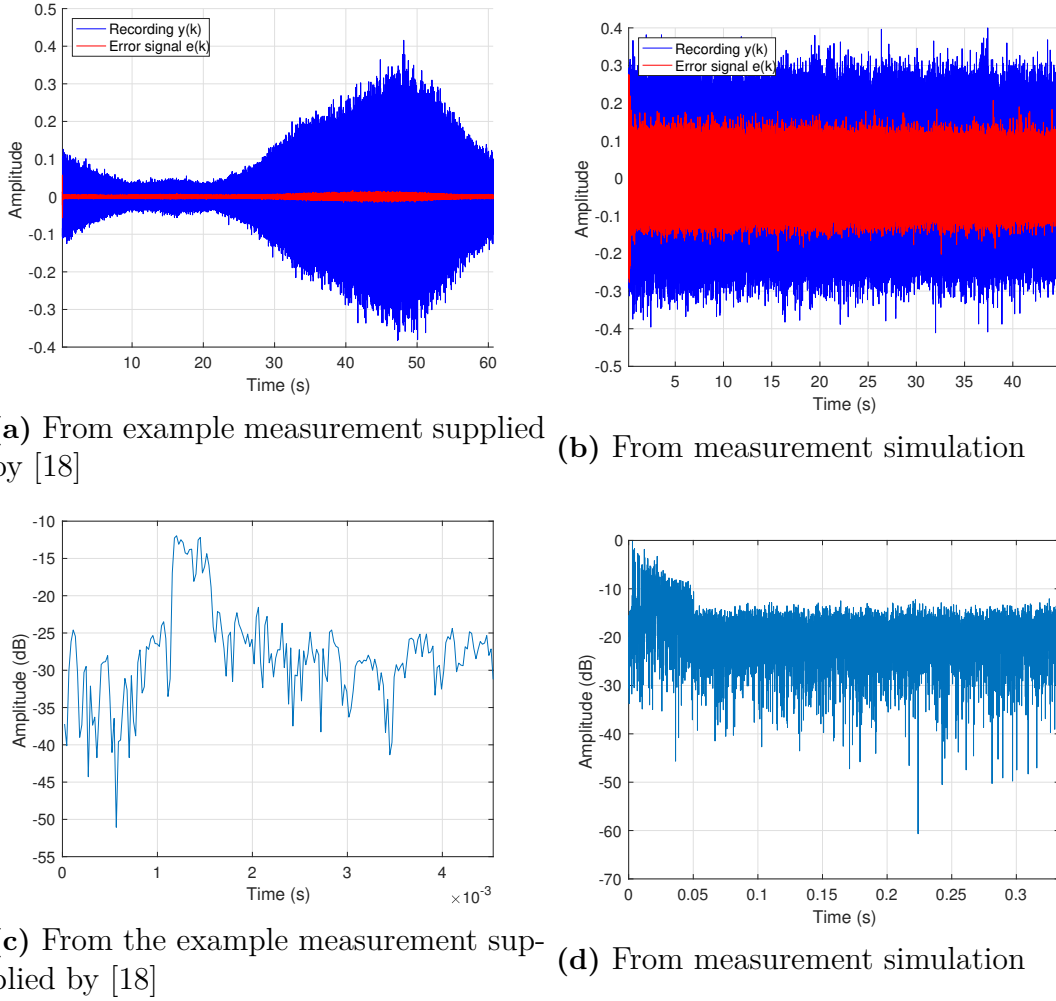


Figure 3.2: Comparison: measured signal and error signal, respective IR and backward integration

A solution to this problem in processing the estimated RIRs and extending them beyond their noise floor is suggested by N. Bryan and J. Abel in [19]. There are two methods proposed, the one used in this thesis is the *Natural Extension Synthesis* which preserves the natural noise sequence found in the original RIR by windowing its late-field.

The process of this technique consist of splitting the RIR $h(t)$ into frequency bands $h_k(t)$. Frequency-dependent energy profiles $\hat{\beta}_k(t)$ are then computed by smoothing the square of $h_k(t)$ over a running window. Initially, in this project, the responses were not split in frequency bands to simplify the process. As a consequence, the extended RIRs did not recover the true responses and the results of the parameters returned implausible data. After the identification of this problem, the split of the responses was introduced using a third-octave filter bank. This did not present a perfect magnitude reconstruction, but it is considered acceptable. The magnitude transfer function of the filters are presented in Figure 3.3.

To compute the window λ_k , every frequency band energy profile is modelled as an exponential decay plus a noise floor, according to Eq. 3.5.

$$\beta_k(t; \Theta) = \sigma_k^2 + \gamma_k^2 e^{(-2t/\tau_k)}, \quad (3.5)$$

Where Θ is a vector containing three parameters: $\Theta = [\sigma^2 \gamma^2 \tau]^T$. The three parameters correspond to the noise floor variance σ^2 , the initial late-field gain γ and the late-field time constant τ .

The parameters of Θ are estimated simultaneously via a sliding least-squares estimate technique on the decibel level of the smoothed energy profiles. With the obtained Θ and λ_k , the extended single frequency band responses $h_k(t)$ are computed and the final extended RIR is reconstructed.

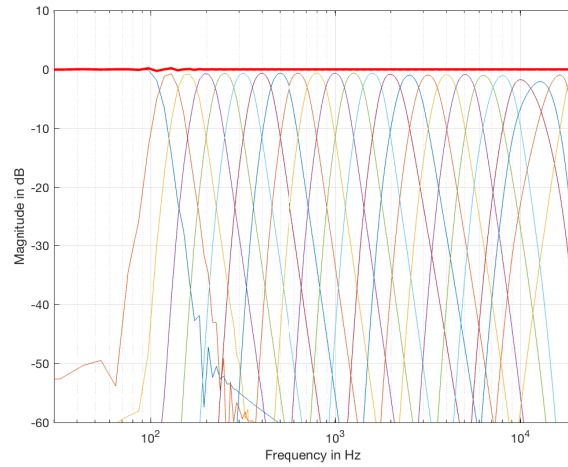


Figure 3.3: Filter Bank Transfer Function Magnitudes

The spectrogram of the RIR before and after the processing, the resulting responses and related backward integration curve are presented in Figure 3.4.

3.4 Computation of the Objective Parameters

Most of the objective parameters are obtained through the use of the RIR, but in some cases it is necessary to have assumptions. This section will present the assumptions considered for the calculation of the interaural cross correlation and the Reverberation time.

3.4.1 Interaural Cross Correlation Assumptions

For the determination of the inter-aural cross correlation, a binaural measurement is required. To perform a binaural measurement, a special setup is required, with a *dummy head* or a simplified alternative like the *Jecklin disc*. Unfortunately, these devices are too heavy to be mounted at the external end of the moving arm used for the measurement. To solve this problem an investigation has been carried out to find a possible alternative of using only a pair of microphones. The starting point of this investigation is the correlation between two omnidirectional microphones

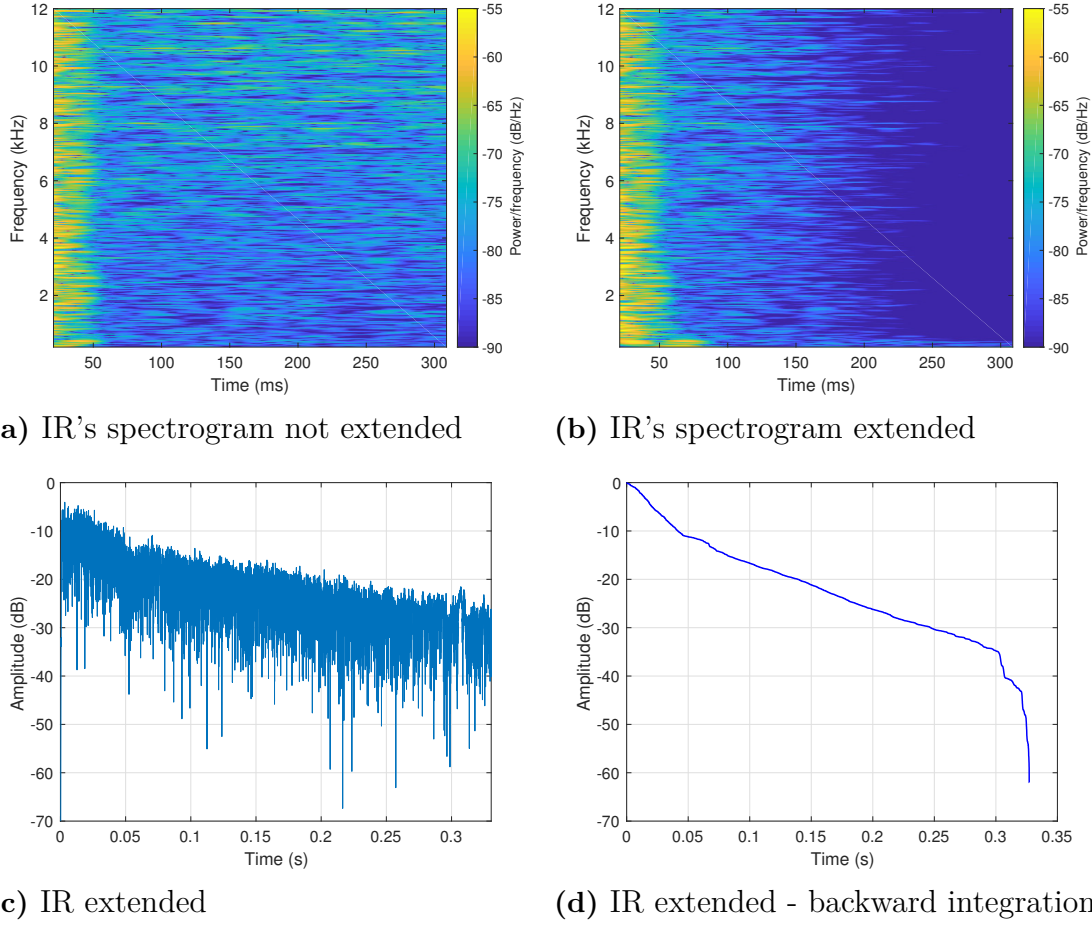


Figure 3.4: Result of the process of the extend IR

spaced equally to the distance separating the ears on a human head (approximately 17 cm) and the same couple of microphones with an interposed head as presented by Lindevand and Benade in [20] —showing that the theoretical correlation of two omnidirectional microphones separated by a distance d will correspond to the behaviour of a sinc function:

$$R(kd) = \frac{\sin(kd)}{kd} \equiv \text{sinc}(kd), \quad (3.6)$$

where $k = \omega/c$. In the paper it is shown that the correlation of the same setup with the addition of an interposed head corresponds to a modified sinc function:

$$R(kd)_{with_head} = \frac{\sin(\alpha kd)}{\sqrt{1 + (\beta kd)^4}}, \quad (3.7)$$

where $\alpha = 2.2$ and $\beta = 0.5$.

Considering Eq. 3.7, different simulations have been done to determine whether it is possible to obtain data of magnitude squared coherence (MSC) closer to the ones resulting with a dummy head measurement with just a simple pair of microphones. The simulations consist of a binaural recording, in a diffuse field, with two

microphones of the same pick up pattern at a distance of 17 cm from each other, oriented in opposite directions (0° and 180°), without any interposing. From the data acquired, the MSC has been calculated and compared with the different results from the different microphones pairs. The MSC of a human related transfer function retrieved from the Center for Image Processing and Integrated Computing (CIPIC) HRTF Database has been used as a further reference for the test. The results are presented in Figure 3.5 where it is possible to deduce that the set up with omnidirectional microphones have the data of MSC closer to the theoretical ones. This result is further supported in research carried out by Brandstein and Ward in [21] where similar results were obtained (see Figure 3.6).

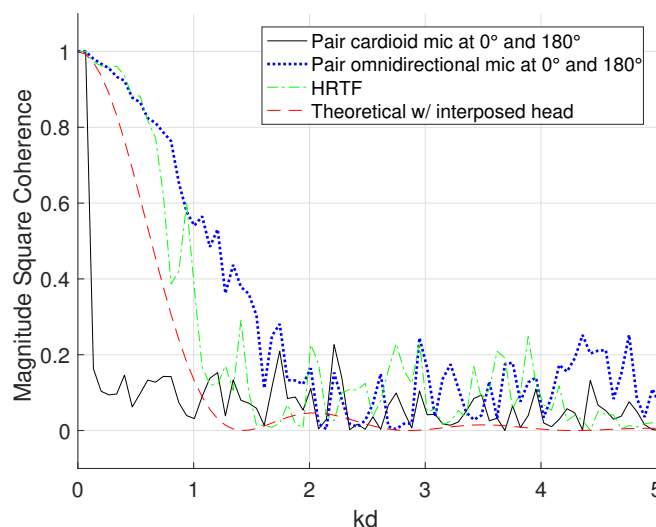


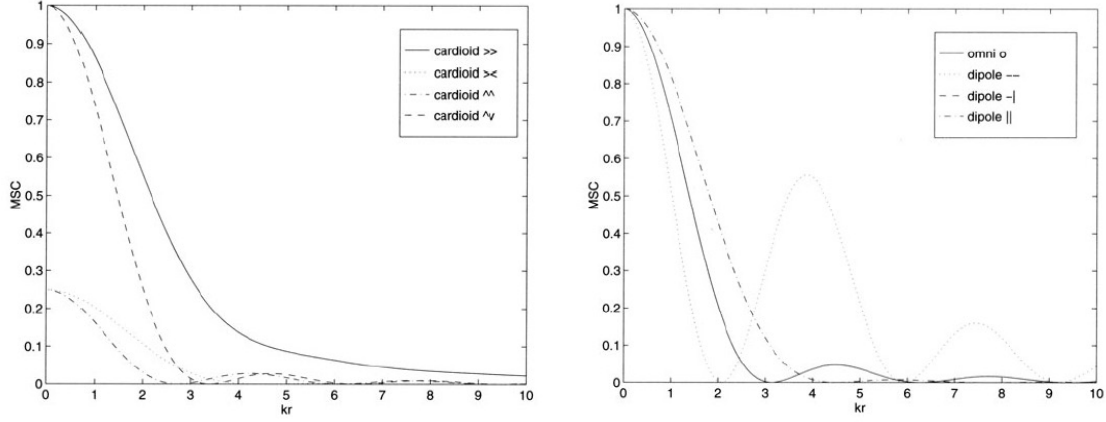
Figure 3.5: Magnitude squared coherence of different set up of a pair of microphones at 17 cm distance from each other

A further test has been done in a simulated environment, recreating the experimental measurement similar to the one done by Käsbaach *et al.* in [22]. The test consists of a binaural recording in an anechoic environment using the same HRTF from CIPIC, as previously mentioned and the two omnidirectional microphones. These were set up to be oriented in opposite directions (0° and 180°).

The receivers are positioned at 1.8 meters distance from the source, in front of the imaginary face. The source consists of a standard stereo set up (two sources at 30° on left and on the right sides). The stimuli consist of three band-limited noise signals with centre frequencies at 250 Hz, 1 kHz and 4 kHz with bandwidth of 2 octaves and a high pass filtered noise signal with cut-off frequency at 8 kHz. Both filters are digital Butterworth, of the eighth and fourth order respectively. The signals are generated to obtain an inter-channel correlation of 0, 0.3, 0.6, 0.8 and 1. The result of the measurements from the two set ups are compared in Figure 3.7.

The data of these plots can be considered acceptable considering the statement from Blauert and Lindemann in [23] about the perception of the spatial extent: their research shows that the perception of the split into two components starts under a

correlation of 0.4.



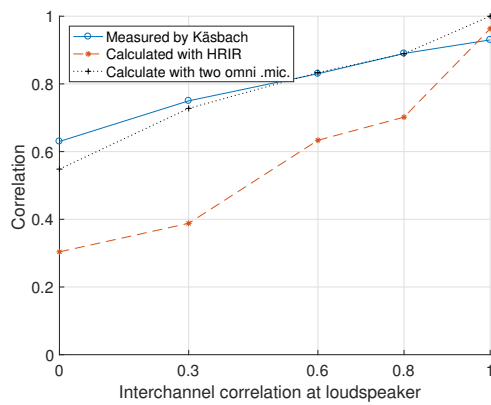
(a) Magnitude squared coherence for various orientations of cardioid microphones in a spherical isotropic noise field (b) Magnitude squared coherence for omnidirectional and dipole microphones in a spherical isotropic noise field

Figure 3.6: Magnitude squared coherence for different microphones according to Brandstein [21]

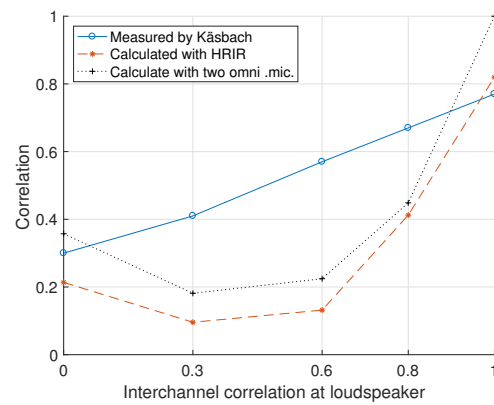
To carry out the binaural measurements, only one microphone has been used, taking advantage of the spatial resolution of the positions estimated. In this way the measured IR of the second microphone correspond to the IR estimated on a position, with a distance corresponding to the distance between the ears of a human head (approximately 17 cm). Considering that the radius of the circular trajectory used is of 1 meter, the interval of 10° , which correspond to a distance of 17.45 cm, has been used.

3.4.2 Reverberation Time Computation

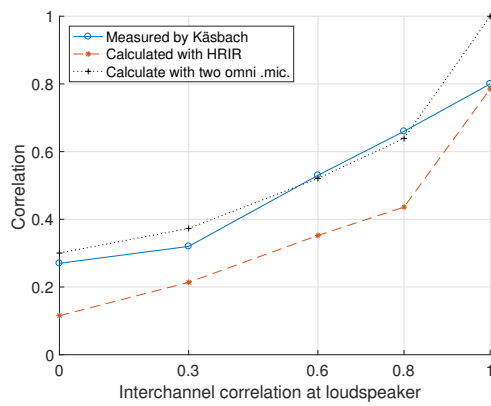
When calculating the reverberation time, as mentioned by Kuttruff in [8], the upper limit of the integral in Eq. 2.2 must be limited to a finite value, discarding the noise portion from the IR. This limit corresponds to the point in time when the decaying of the excitation reaches the noise floor level. When calculating the RT with the *extended* IRs, it is no longer necessary to truncate the IR because, after its extension (Sec. 3.3), the noise floor is eliminated. Considering this, the calculation is carried out by a process which iterates on all the IRs of the measurement. It consists of determining automatically the two points on the backward integration curve where its amplitude corresponds to both -5 dB and -35 dB (for the case of T_{30}). Then using all the data within the two points of the Schroeder curve, it uses the least mean square method to fit a line and finally determining the $RT = \frac{60}{|d|}$ where d is the slope of the line.



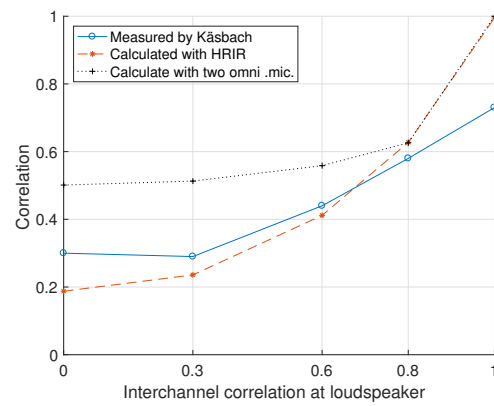
(a) Band limited noise = 250 Hz



(b) Band limited noise = 1 kHz



(c) Band limited noise = 4 kHz



(d) Band limited noise = 8 kHz

Figure 3.7: Correlation for two omnidirectional microphones and HRIR with band limited noise

3.5 Measurements

The room acoustic measurements took place at the division of Applied Acoustic at Chalmers University of Technology, in the reverberation chamber, the Audio lab and listening room. This way it allows to compare the parameters of three different environments. Three different measurements have been carried out, a brief overview is presented in Table 3.1 and described in detail in the following sections.

Table 3.1: Measurements overview

	Meas. Type	Source Type	Mic. Type	Position	Purpose
1	Moving microphone	Omnidirectional	Omnidir.	Centre & Corner	Compare the averaged data to the standard results
			B-Format		
2	Moving microphone	Directional	Omnidir.	Centre & Corner	Study the local data variations
			B-Format		
3	Static microphone	Directional	Omnidir.	Centre	Compare the results of the two systems
			B-Format		

3.5.1 Moving Microphone Measurements

Two set-ups have been used: one for the measurement of the parameters inspired by the standard ISO 3382-1 and the other with a directional source to simulate the case of a class room or conference room.

The devices used for the set-up of the measurement according to the standard are:

- First order B-format microphone Sennheiser AMBEO VR Mic;
- Electret omnidirectional microphone with related preamplifier;
- Variable spherical scanning array system VariSphear (Figure 3.9),
- Sound interface PreSonus firestudio project
- Script for the acquisition of the measurement: Acoustics Hardware - Python script for advanced measurement by Carl Andersson (<http://acoustics-hardware.readthedocs.io/en/latest/>)
- Omnidirectional loudspeaker B&K *omnisource 4295*;
- Power amplifier B&K *type 2735*.

The devices used for the setup with directional source are:

- First order B-format microphone Core Sound TetraMic S/N 2260 (Figure 3.8) with the related four preamplifiers,
- Electret omnidirectional microphone with related preamplifier;
- Variable spherical scanning array system VariSphear (Figure 3.9),

- Sound interface PreSonus firestudio project
- Script for the acquisition of the measurement: Acoustics Hardware - Python script for advanced measurement by Carl Andersson (<http://acoustics-hardware.readthedocs.io/en/latest/>)
- Genelec active loudspeaker model 8020B.



Figure 3.8: B-Format microphone - **Figure 3.9:** VariSphear - [24]
Core Sound TetraMic

The measurements have been carried out at two positions in every room: one close to the corner and one approximately in the centre of the room, Figure 3.11 explains in detail the positions of the VariSphear and the loudspeaker in the rooms. The radius of the circle corresponding to the trajectory of the microphone is of one meter. The internal parameters used in the programming of the VariSphear are: *target acceleration* = 600, *target velocity* = 360/ revolution time. The revolution time used is 9 minutes, plus 15 seconds before the arm starts to move, with an additional 15 seconds at the end of the revolution. The additional 30 seconds will be discarded by the Enzner's estimation function through the parameters to define the overlap measurement. The first 15 seconds are considered to excite the room to a steady state.

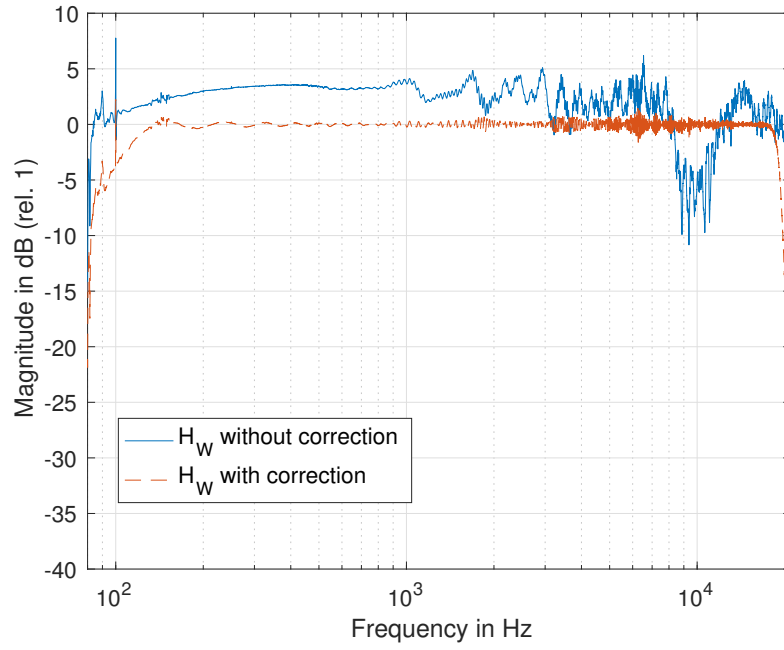
The long revolution time is a consequence of the tests done in the simulation process. While in [7] the measured signal and the estimated error is obtained with a revolution time of one minute, in an anechoic environment, it has been noticed that in the case of a longer reverberation time it is necessary to have a slower movement to achieve a small estimation error.

The acceleration of the VariSphear is set to a high value to permit the device to reach a constant velocity immediately after the start of the revolution. In the case of low acceleration, especially with shorter revolution time (i.e. higher revolution velocity), the VariSphear will constantly accelerate, until it reaches the set velocity and then it will decelerate, approaching to the stop point.

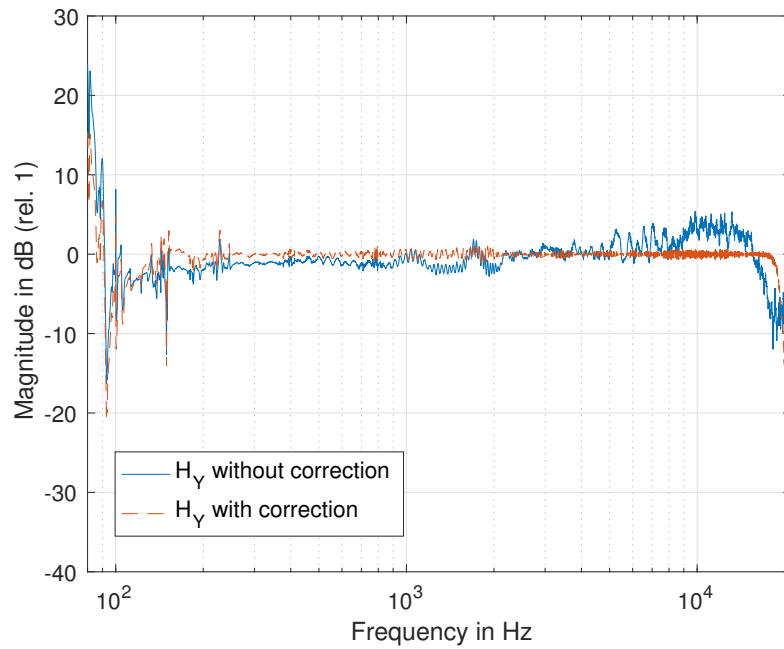
The start position (0°) of the measurement is oriented according to Figure 3.11, the direction of rotation of the arm is counterclockwise.

After the acquisition, the data from the B-format microphone have been processed to obtain the omnidirectional signal and the figure of eighth signal. This is calculated with the microphone constantly oriented in the same direction independent of the arm position. The space interval for every IR was set to one degree, which corresponds to 1.74 cm. The data of the B-format measurements have been filtered to correct the alteration of the equalisation response of the microphone when converting the data from the A-format to the B-format. Then the data is calibrated to maintain the same level for all the four microphones. Figure 3.10 from [25] shows the response of the microphone with and without the filter.

The data acquired with the Core Sound TetraMic, before it is filtered with the aforementioned filter, has been processed with a calibration filter created with a reference microphone for every single capsule of the TetraMic to obtain a flat response for every capsule. On every filtering process, the generated group delay is calculated and compensated in the resulting data.

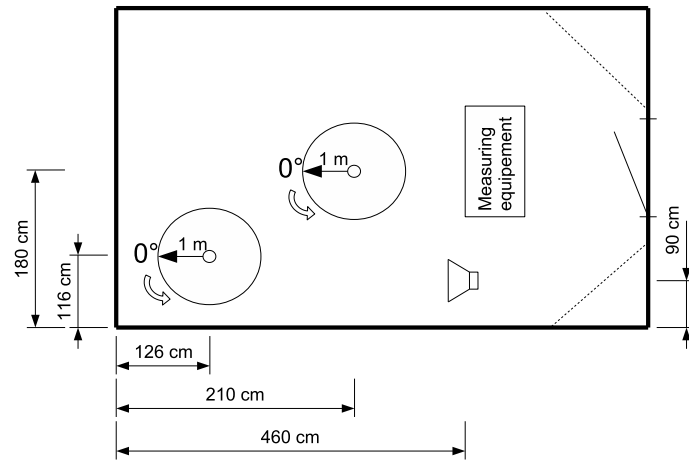


(a) Omnidirectional pattern

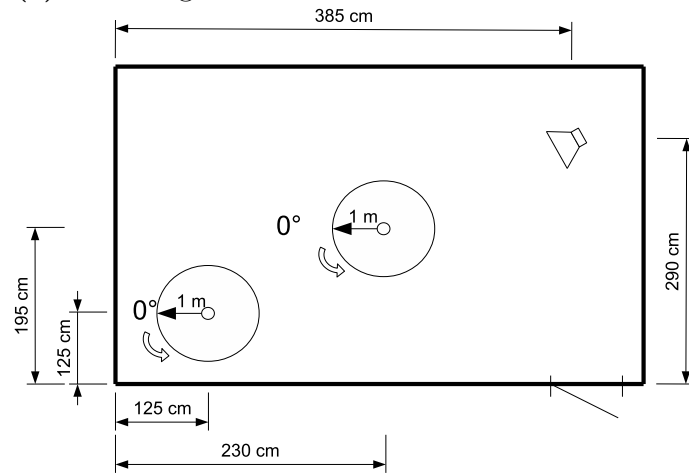


(b) Figure of eight pattern

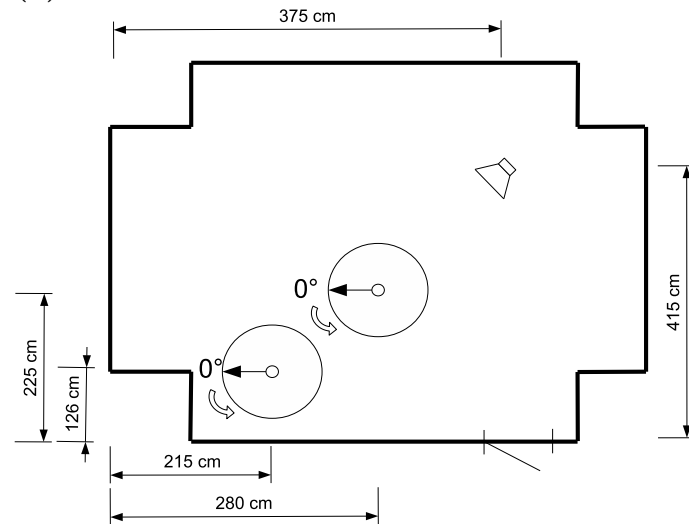
Figure 3.10: Comparison of the filtered and unfiltered B-format microphone [25].



(a) Listening room



(b) Audio Lab



(c) Reverberation Chamber

Figure 3.11: Position of the VariSphear and loudspeaker in the different Labs.

3.5.2 Static Measurements - Directional Source

Measurements with a static microphone have been carried out in four positions on the trajectory of the moving microphone, with the aim to compare the results of the two systems. The four positions are at 0° , 90° , 180° and 270° in the centre of the Audio lab and Reverberation chamber. In addition to these four positions, as a consequence of the behaviour observed in the Listening room, the measurements have been taken in two other positions corresponding to the one in front of the loudspeaker and the one just off the front of the loudspeaker. These will correspond to 213° and 227° in the Audio Lab and; 247° and 256° in the Reverberation chamber. This could not been done in the listening room, as it was not available at that time. The devices used in the setup of the measurements are:

- First order B-format microphone Core Sound TetraMic S/N 2260 (Figure 3.8) with the related four preamplifier,
- Electret omnidirectional microphone with related preamplifier;
- Variable spherical scanning array system VariSphear (Figure 3.9),
- Sound interface PreSonus firestudio project
- Script for the acquisition of the measurement: Acoustics Hardware - Python script for advanced measurement by Carl Andersson (<http://acoustics-hardware.readthedocs.io/en/latest/>)
- Genelec active loudspeaker model 8020B.

Before the impulse responses were estimated, the data of the B-format microphone have been filtered as previously shown in Sec. 3.5. The stimuli signal used is a white noise, the duration of the recording is one minute and the estimation of the impulse responses has been computed with the H1 estimator. The computation of the parameters have been carried out with the same process used with the moving microphone, with exception of the reverberation time. The MATLAB function *calT60ters* implemented by Peter Svensson, supplied by the division of Applied Acoustic of Chalmers University of Technology, has been use in this case. The script has been re-adapted to compute the RT in octave bands instead of third octave bands.

3.6 Perceptual Evaluation

An informal listening evaluation was taken in order to validate the results of room acoustic parameters. The experiment consisted of the convolution of different IRs estimated by the system, with anechoic recordings. The results were played using the built-in sound card of a Macbook pro, headphones Beyerdynamic DT 770 PRO and AKG K702.

The anechoic stimuli where made from a male singing voice recorded at the Cologne University of Applied Sciences, available at <http://audiogroup.web.th-koeln.de/anechoic.html>. Only two people took part in the experiment, both reported of no having any hearing impairments and had experience in evaluation tests. One of them is considered to be an expert in the auditory event field.

The positions of the convolved IRs have been chosen considering the same room acoustic parameters: positions with different value at relevant distances and posi-

3. Methods

tions with different value at very short distance. Table 3.2 presents the positions and relative parameters used in the test.

Table 3.2: Positions of the IR used in the perceptual evaluation test.

Room	Parameter	Parameter Value	Position
Listening room centre	C_{50}	21.4 dB	173°
Listening room centre	C_{50}	17.1 dB	185°
Listening room centre	C_{50}	11.6 dB	309°
Listening room centre	T_s	33 ms	150°
Listening room centre	T_s	26 ms	147°
Listening room centre	$IACC$	0.40	66°
Listening room centre	$IACC$	0.63	71°
Listening room centre	$IACC$	0.43	139°
Listening room centre	$IACC$	0.85	159°
Listening room centre	T_{30}	0.13 s	143°
Listening room centre	T_{30}	0.50 s	93°
Listening room centre	T_{30}	0.40 s	121°
Audio Lab centre	C_{50}	8.1 dB	24°
Audio Lab centre	C_{50}	3.2 dB	47°
Audio Lab centre	STI	0.60	280°
Audio Lab centre	STI	0.77	219°
Audio Lab centre	T_s	57 ms	49°
Audio Lab centre	T_s	46 ms	39°
Audio Lab centre	$IACC$	0.59	198°
Audio Lab centre	$IACC$	0.30	213°
Audio Lab centre	$IACC$	0.70	226°
Audio Lab centre	EDT	0.80 s	326°
Audio Lab centre	EDT	0.50 s	343°
Audio Lab centre	T_{30}	1.06 s	75°
Audio Lab centre	T_{30}	0.80 s	86°
Audio Lab centre	T_{30}	1.10 s	335°
Audio Lab centre	T_{30}	0.80 s	340°
Reverberation chamber centre	C_{50}	-4.0 dB	134°
Reverberation chamber centre	C_{50}	-6.0 dB	153°
Reverberation chamber centre	C_{50}	-2.0 dB	247°
Reverberation chamber centre	STI	0.46	265°
Reverberation chamber centre	STI	0.37	281°
Reverberation chamber centre	T_{30}	2.79 s	67°
Reverberation chamber centre	T_{30}	2.30 s	71°
Reverberation chamber centre	EDT	2.54 s	175°
Reverberation chamber centre	EDT	2.30 s	166°
Reverberation chamber centre	EDT	2.25 s	187°

4

Results

In this chapter the results from the measurements are showed. The chapter is divided into four sections. The first section presents the averaged data of the room acoustic parameters, computed in octave bands, from the measurements taken with the omnidirectional source. For a comparative purpose, the plots show also the results of a standard measurement done previously at the division. The second section displays the data resulting from the estimated IRs of the moving microphone in the different positions over the trajectory. The third section represents the comparison of the parameters' data, measured in the same positions with a moving microphone and a static microphone. The fourth section reports the results of the informal perceptual evaluation.

The STI result are presented in octave bands. Although the modulated transfer index (MTI) is showed instead the STI. To avoid confusion, the captions of these figures will contain the nomenclature STI instead of MTI even if it is incorrect.

4.1 Averaged Results - Omnidirectional Sources

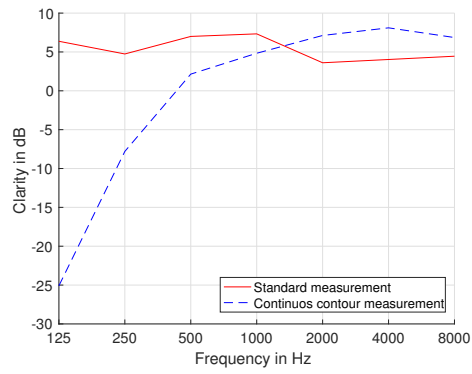
The results of a standard measurement, previously done at the Applied Acoustic division of Chalmers University, are compared with the results of the system studied in this project to verify if the data obtained are compatible.

The parameters of the Audio Lab are presented in Figure 4.1 and the ones of the Reverberation chamber in Figure 4.2.

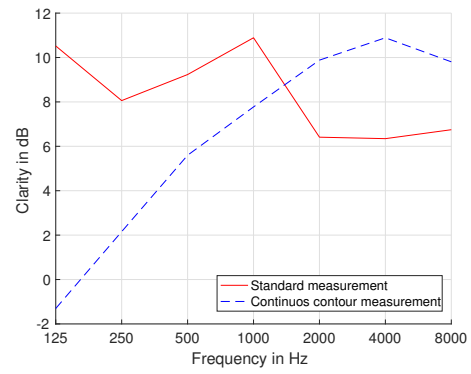
It must be considered that the process used to extend the IRs beyond its noise floor could have introduced some alteration to the final IRs. However, considering the just noticeable difference of the parameters presented in Sec. 2.6.1, the curves in the plots present small variations, giving a validation to the results obtained with the moving microphone system. Considering that the system return 360 measurement positions, it is possible to assert that, for the cases under study, where the room have small dimensions, the results of this system can return more accurate mean data than from a standard measurement with less measurement spots.

The only significant difference which emerge from this comparison is on the clarity and definition data. The plots in Figure 4.1a, 4.1b, 4.1c, 4.1d and 4.2a, 4.2b, 4.2c, 4.2d, show that the moving microphone system returned very low values in the two lowest octave bands of these two parameters.

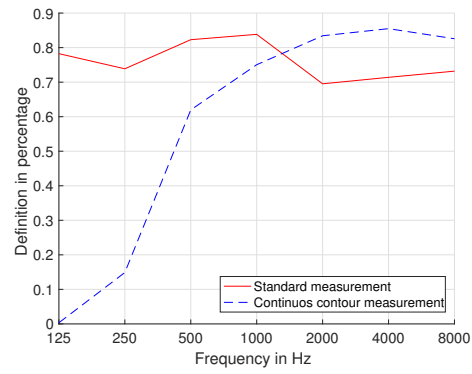
4. Results



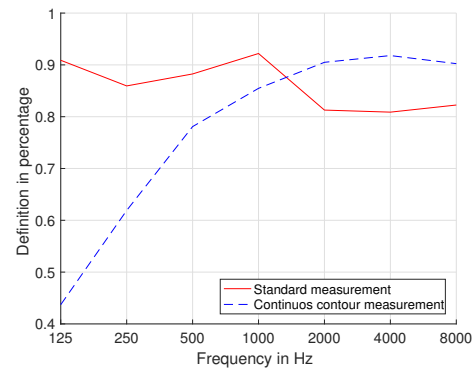
(a) Clarity C_{50}



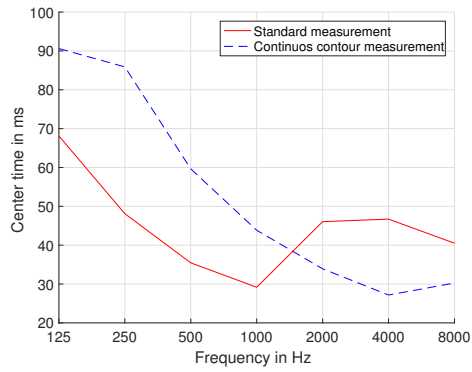
(b) Clarity C_{80}



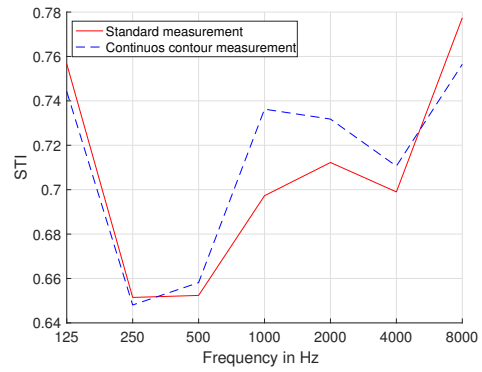
(c) Definition D_{50}



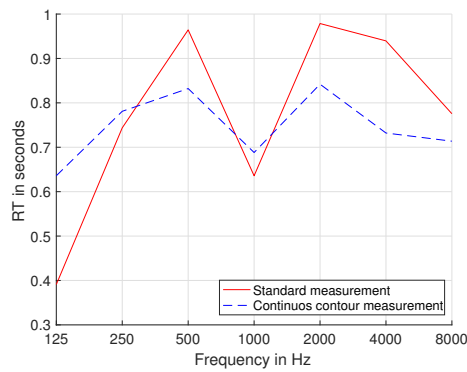
(d) Definition D_{80}



(e) Centre time T_s

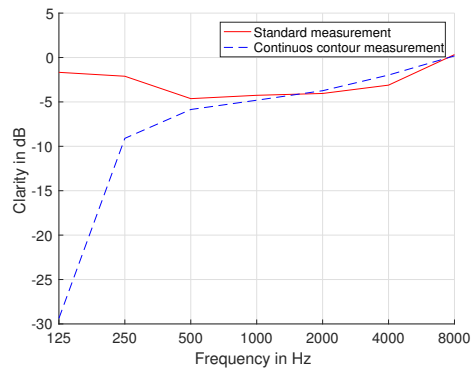
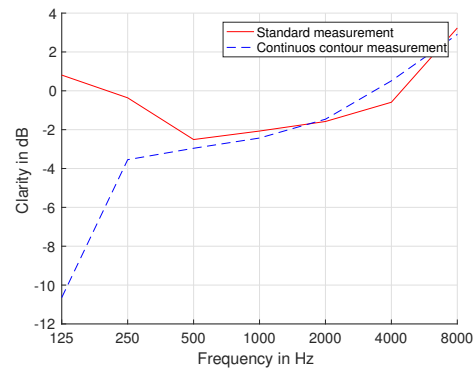
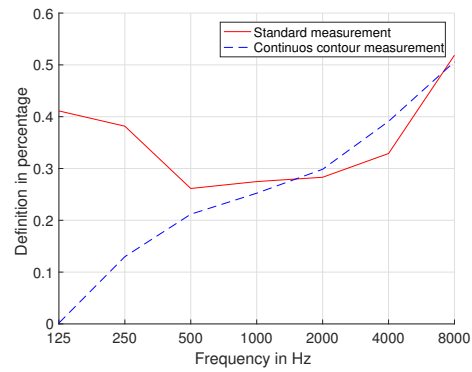
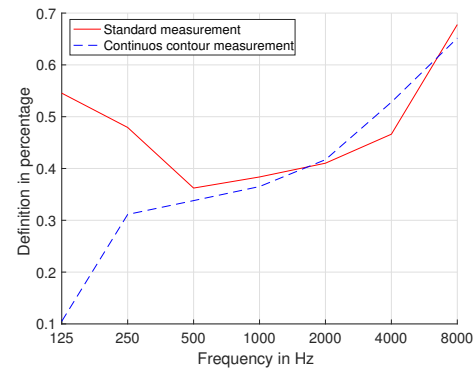
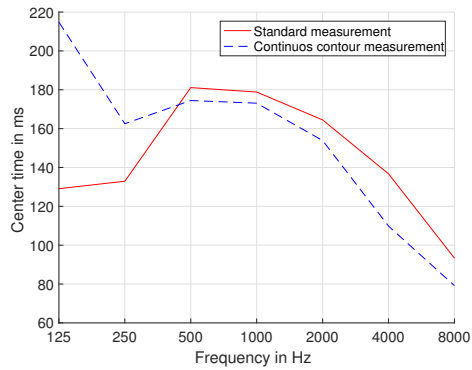
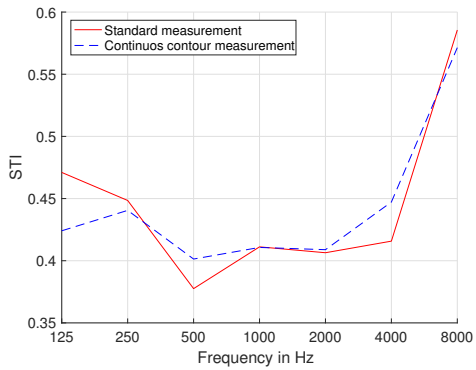
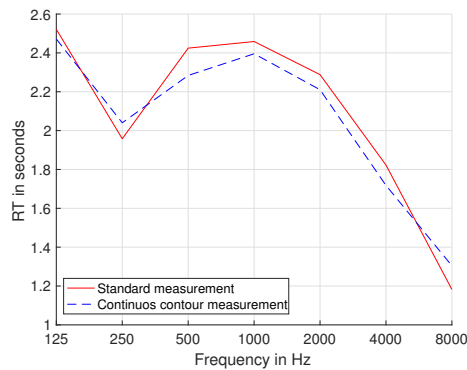


(f) Speech transmission index STI



(g) Reverberation time T_{30}

Figure 4.1: Audio lab parameters comparison


(a) Clarity C_{50}

(b) Clarity C_{80}

(c) Definition D_{50}

(d) Definition D_{80}

(e) Centre time T_s

(f) Speech transmission index STI

(g) Reverberation time T_{30}
Figure 4.2: Reverberant chamber parameters comparison

4.2 Local Parameters - Directional Source

The results in this section show the data of the parameter at each position. The x axis in the plots represent the position, in degrees, of the revolution of the microphone. The *interval resolution* of the measurement positions corresponds to one degree, which is equal to a distance of 1.74 cm between two points. This interval is considered small enough. In the case of higher spatial resolution, like interval of 0.1 degrees, the results show small variations between the different points. These variations are small enough to be considered not too relevant and using a longer interval will return more clearer and readable plots. An example of these two cases are presented in Figure 4.3, where the reverberation time of the Audio Lab is shown with a spatial interval of one degree and 0.1 degree.

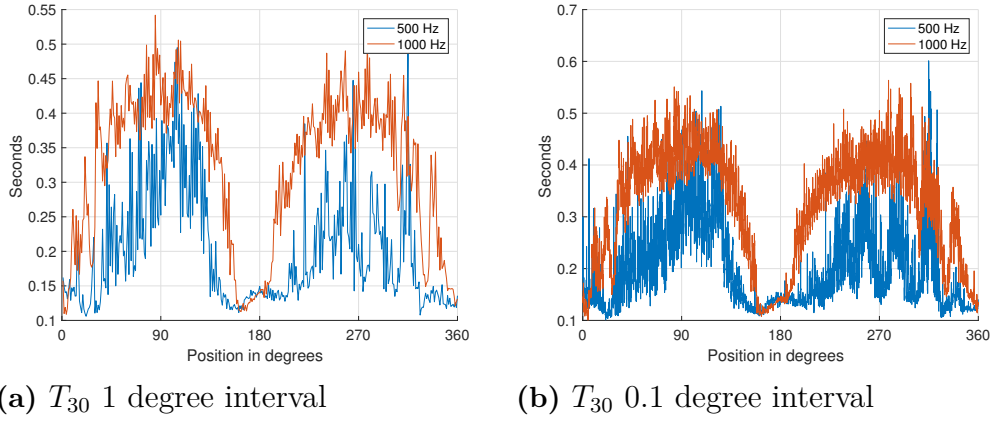


Figure 4.3: Comparison of the result with different resolution over the measurement trajectory - B format mic.

In the following plots (Figures 4.4 - 4.10), only two octave bands are shown: 500 Hz and 1000 Hz because these are considered appropriate and more representative for a source like a human voice. In the figures, the results obtained with the B-format microphone (plot on the left) and with the omnidirectional microphone (plot on the right) are shown.

Comparing the two microphones results, it is possible to observe differences, in almost all the cases, often the discrepancies are minimal and the curves follows similar a trends.

A particular behaviour occurs in the reverberation time data of the Listening room, presented in Figures 4.5a - 4.5d. Here, when the microphone approaches the source, the reverberation time visibly decreases. It has been considered appropriate to carry out two investigations to verify if there is a real change in the reverberation time or if it is a computational error .

The first investigation consists of verifying the Schroeder's curve and the relative computation of the decay time executed by the script on the IRs at two close positions which present a considerable variation in the data: 93° and 143° . Figure 4.6 shows the comparison: the Schroeder's curves are in blue and red are and the lines of the decay computed by the script are in black. The second investigation consists of the perceptual evaluation of the convolution of the two IRs with a anechoic source

(This will be explained in Sec. 4.4). From these investigations, it is possible to attribute this behaviour to the consequence of a calculation error.

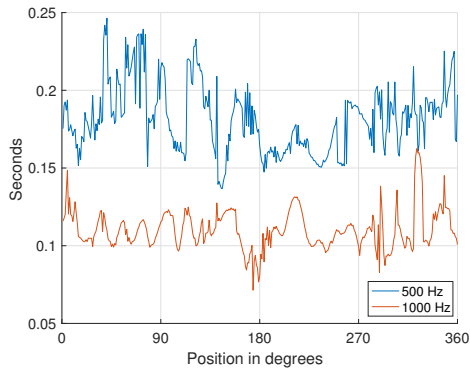
More investigations have been done on the data of the reverberation time, especially in cases where significant variations were obtained in close positions. This last verification shows that the results of the cases under test are calculated correctly (Figure 4.7), but in contrast, it is not perceived as a notable variation in the perceptual evaluation.

The IACC results (Figure 4.12) show clearly when the microphone approaches the source, in all the three rooms. As expected in correspondence to the source, the IACC will have a higher value because at that position the sound wave reaching the listener from the front will produce an almost equal sound pressure at both ears, resulting in a higher correlation value.

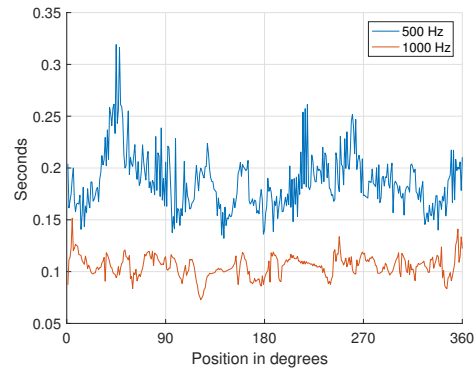
In the case of the measurements in the corner, it is possible to notice the effect of the reflections from the walls, as in Figure 4.12j and 4.12f, where around 90° the reflections increment the IACC value.

Considering the position of the source, it is possible to identify a common behaviour of certain parameters, especially in the measurements taken in the middle of the room. When the position of the microphone approaches the source, the data shows better results (lower value for T_s , higher for C_{50} , D_{50} and STI). This is an expected behaviour, because close to the source, the direct sound will present a higher level than the reflections. It must also be considered that at these positions, the microphone is significantly close to the source. This clarifies the results from the measurements in the corner where this behaviour is not noticeable.

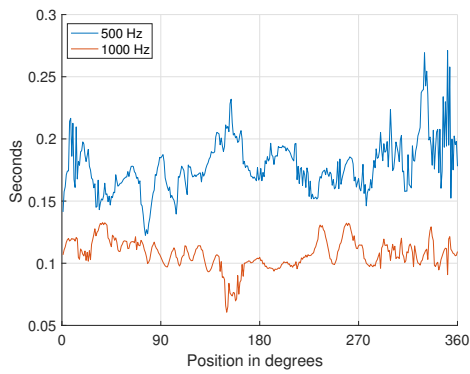
4. Results



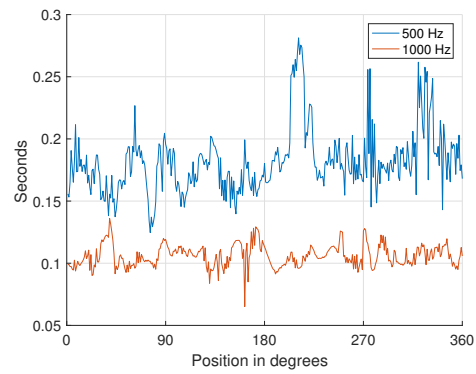
(a) Listening room corner B-format



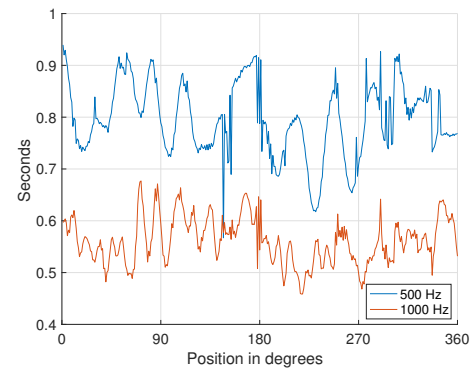
(b) Listening room - corner - omni.



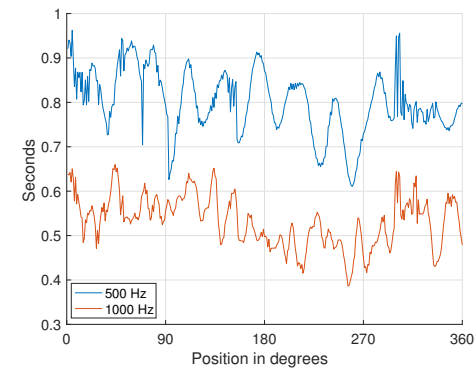
(c) Listening room centre B-format



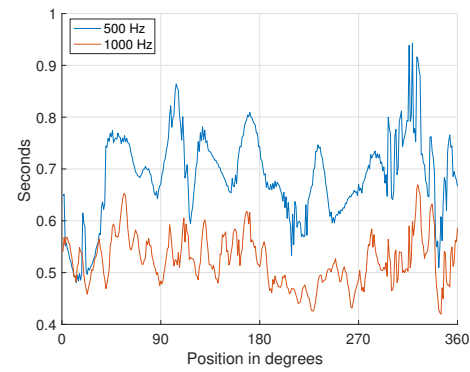
(d) Listening room centre omni.



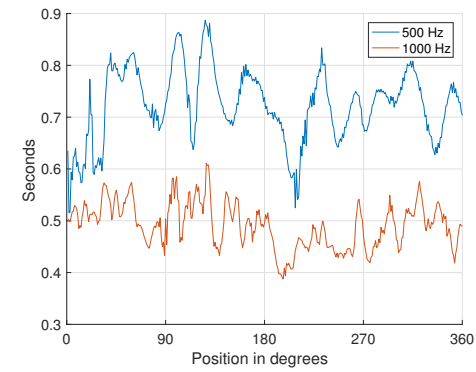
(e) Audio Lab - corner B-format



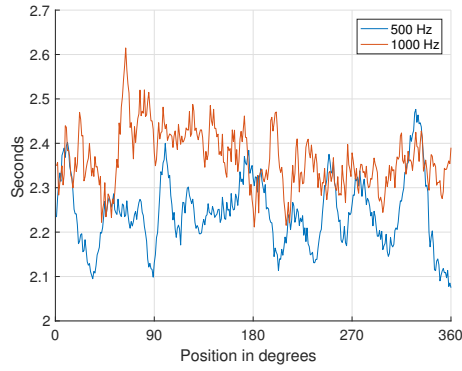
(f) Audio Lab - corner omni



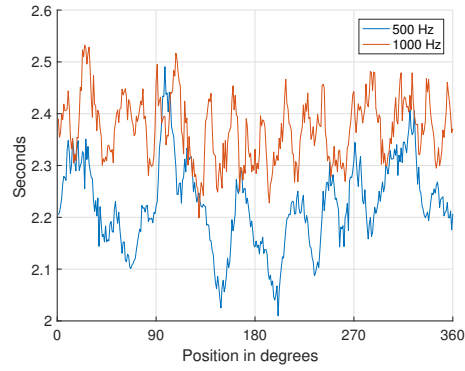
(g) Audio Lab - centre B-format



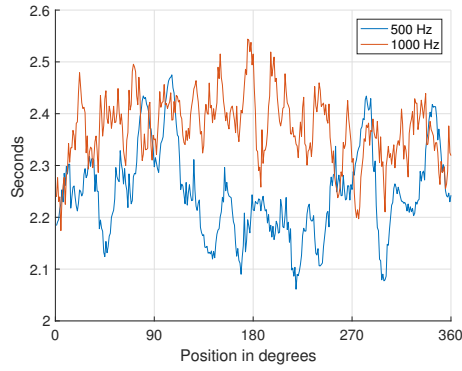
(h) Audio Lab - centre omni



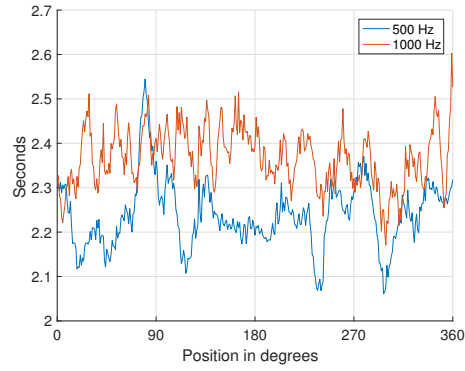
(i) Rev. chamber - corner B-format



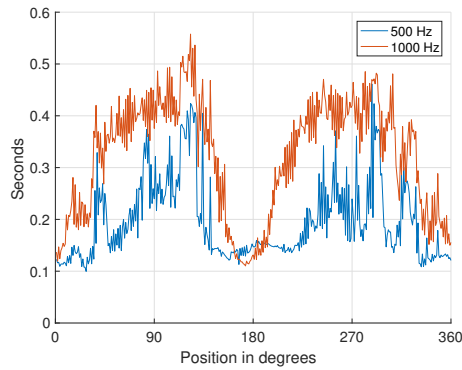
(j) Rev. chamber - corner omni



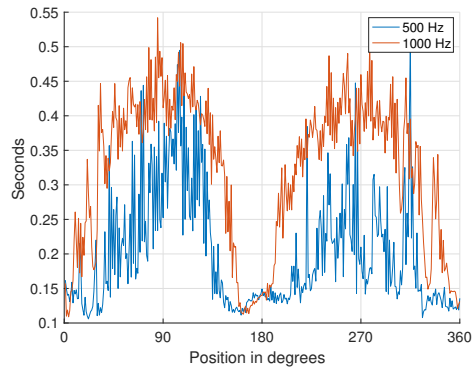
(k) Rev. chamber - centre B-format



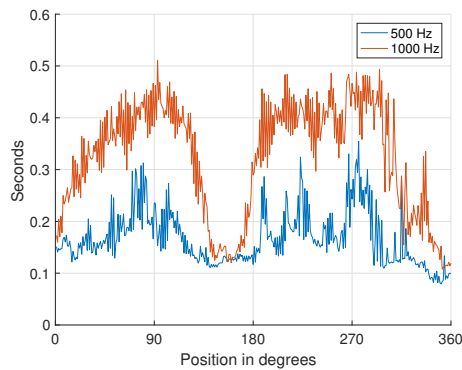
(l) Rev. chamber - centre omni

Figure 4.4: Early decay time - EDT

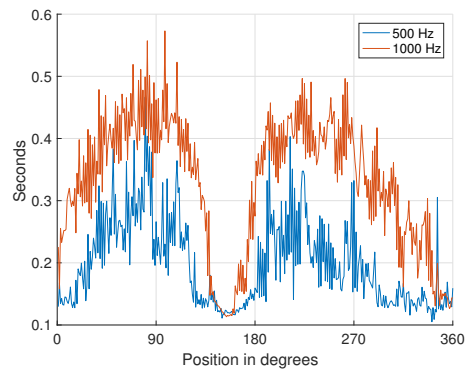
(a) Listening room corner B-format



(b) Listening room - corner - omni.

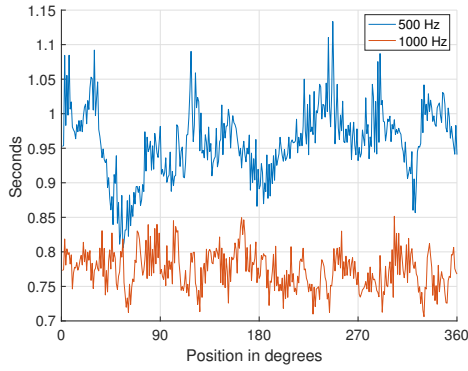


(c) Listening room centre B-format

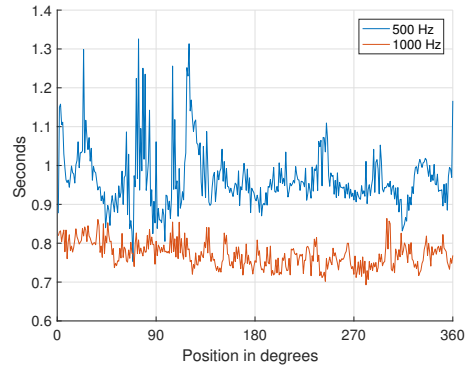


(d) Listening room centre omni.

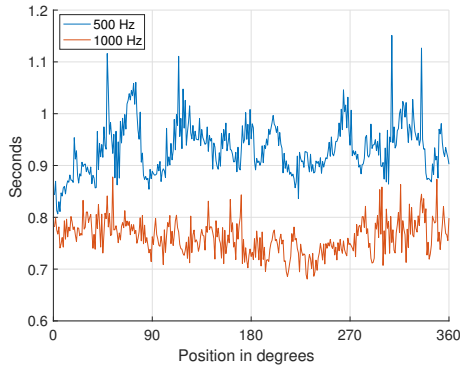
4. Results



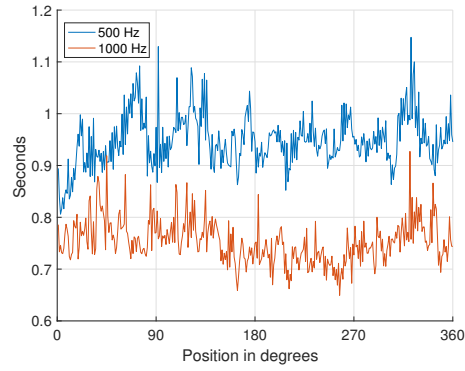
(e) Audio Lab - corner B-format



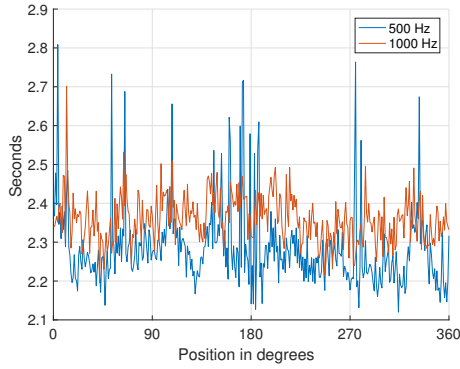
(f) Audio Lab - corner omni



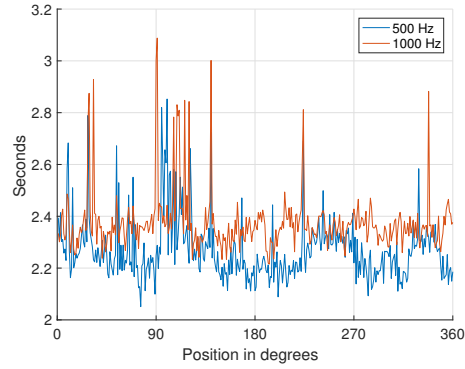
(g) Audio Lab - centre B-format



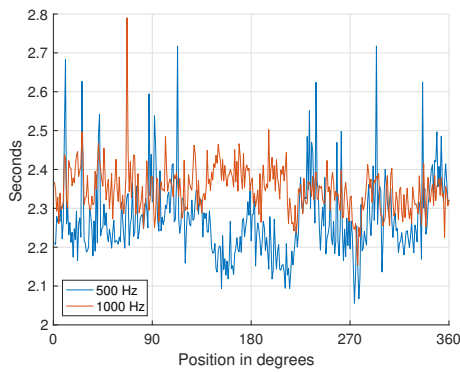
(h) Audio Lab - centre omni



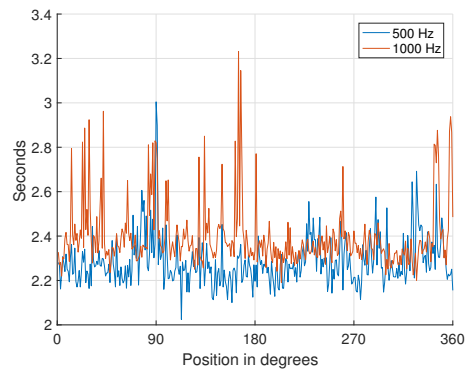
(i) Rev. chamber - corner B-format



(j) Rev. chamber - corner omni



(k) Rev. chamber - centre B-format



(l) Rev. chamber - centre omni

Figure 4.5: Reverberation Time T_{30}

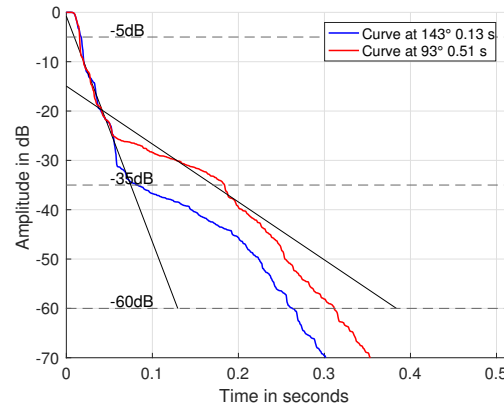
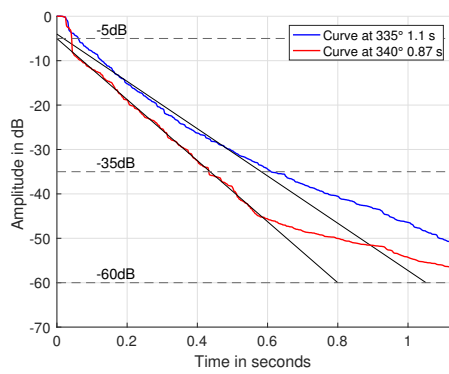
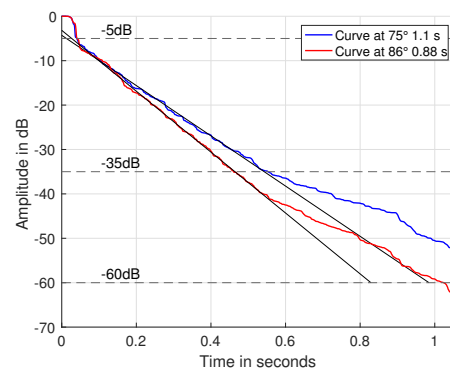


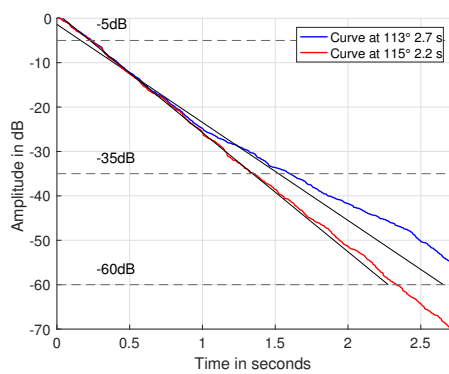
Figure 4.6: Comparison of the Schroeder's curve in the Listening room at 1 kHz



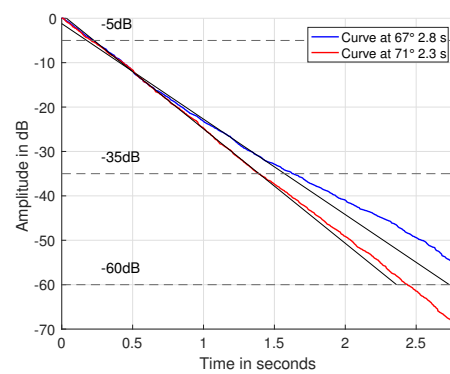
(a) Audio Lab - Curves at 500 Hz



(b) Audio Lab - Curves at 500 Hz



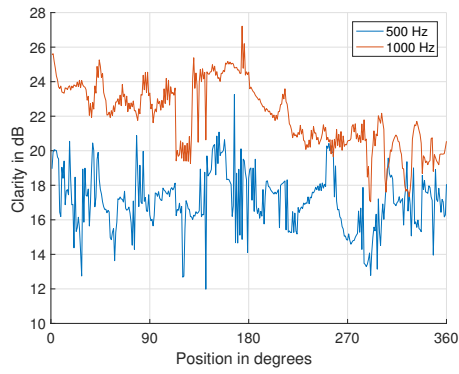
(c) Reverberation chamber - Curves at 500 Hz



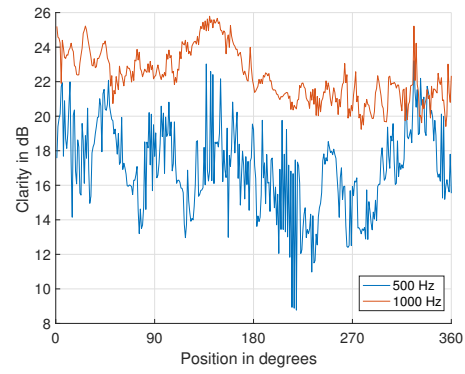
(d) Reverberation chamber - Curves at 1 kHz

Figure 4.7: Comparison of the Schroeder's curve

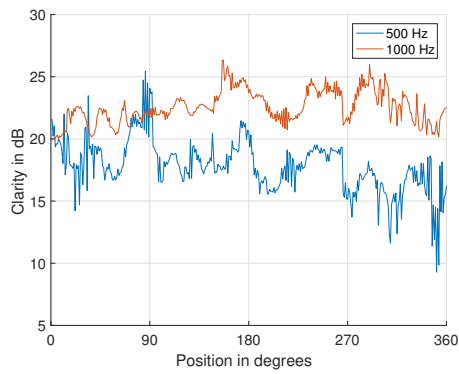
4. Results



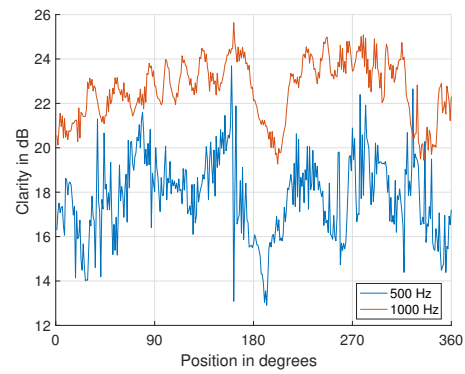
(a) Listening room corner B-format



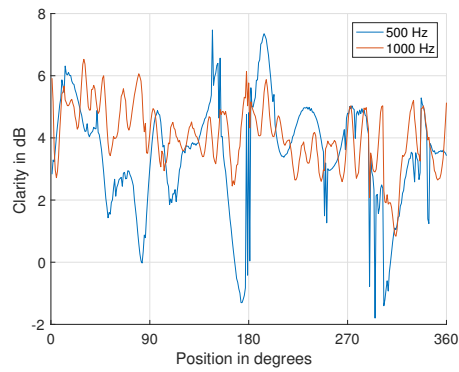
(b) Listening room - corner - omni.



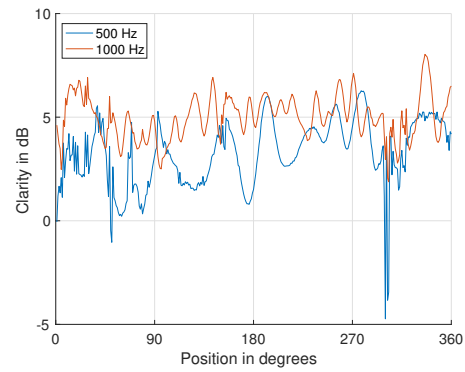
(c) Listening room centre B-format



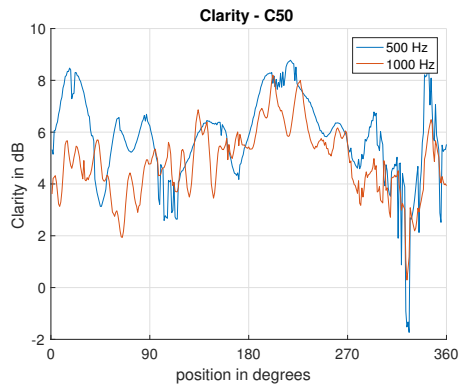
(d) Listening room centre omni.



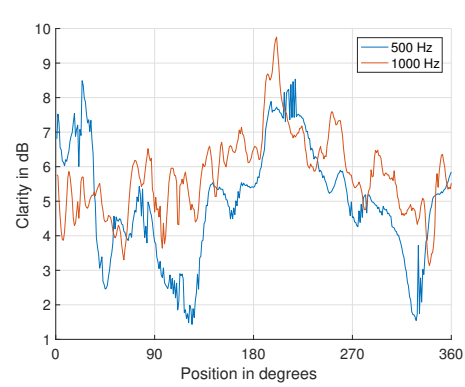
(e) Audio Lab - corner B-format



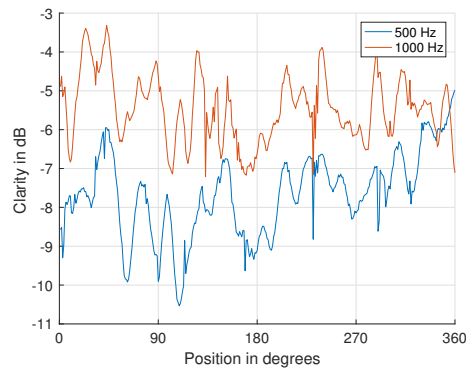
(f) Audio Lab - corner omni



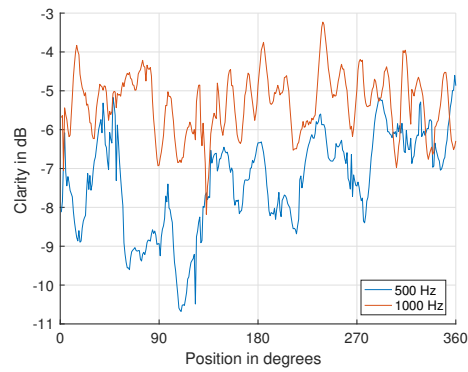
(g) Audio Lab - centre B-format



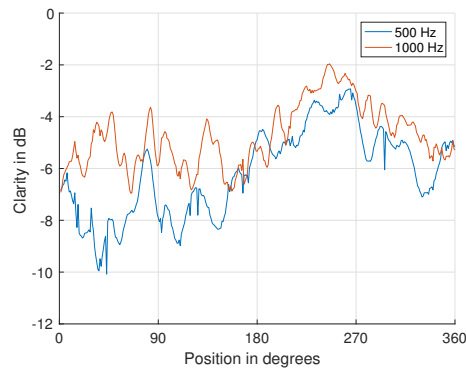
(h) Audio Lab - centre omni



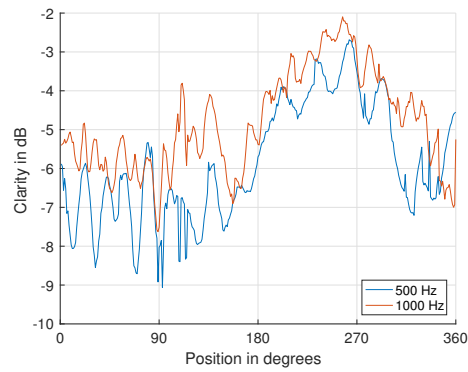
(i) Rev. chamber - corner B-format



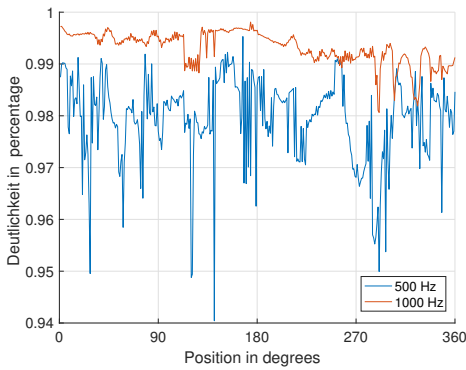
(j) Rev. chamber - corner omni



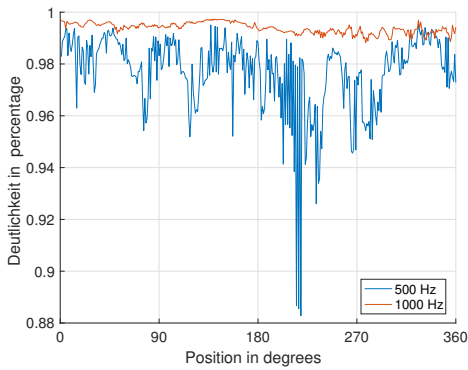
(k) Rev. chamber - centre B-format



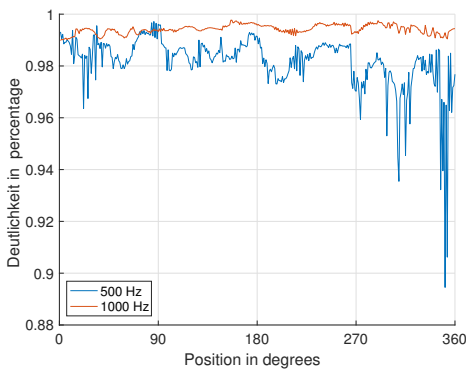
(l) Rev. chamber - centre omni

Figure 4.8: Clarity C_{50} 

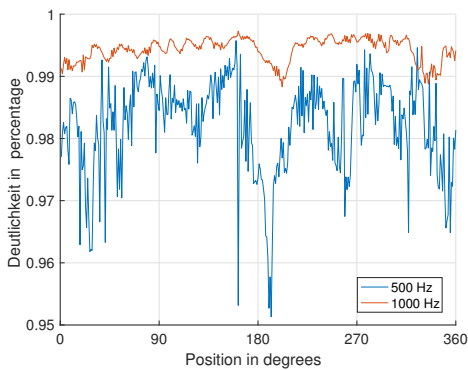
(a) Listening room corner B-format



(b) Listening room - corner - omni.

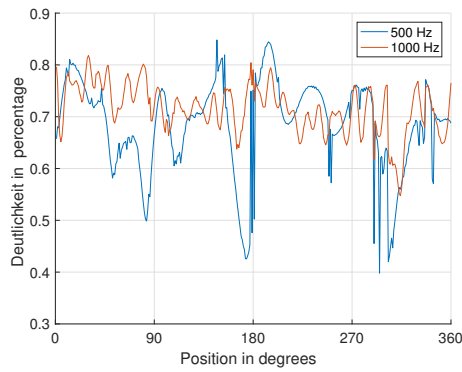


(c) Listening room centre B-format

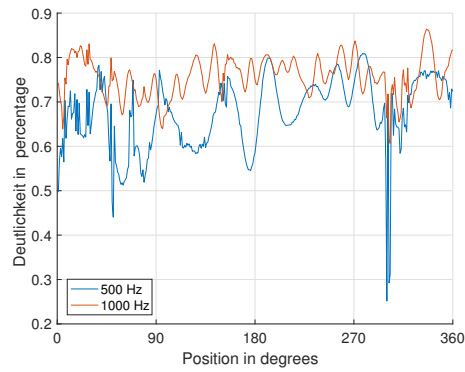


(d) Listening room centre omni.

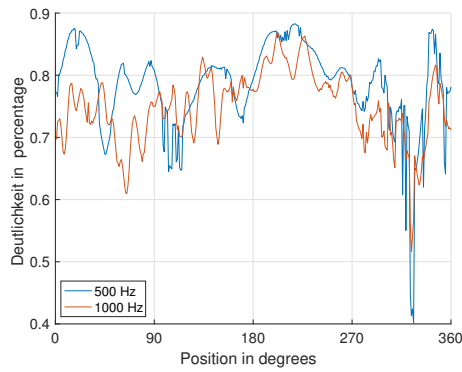
4. Results



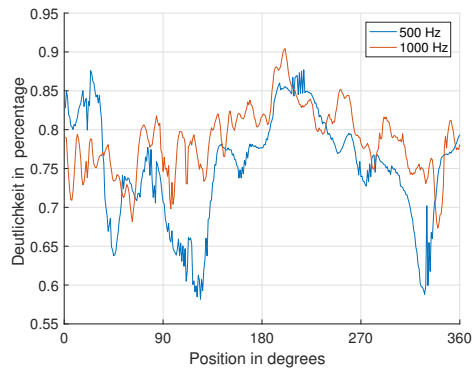
(e) Audio Lab - corner B-format



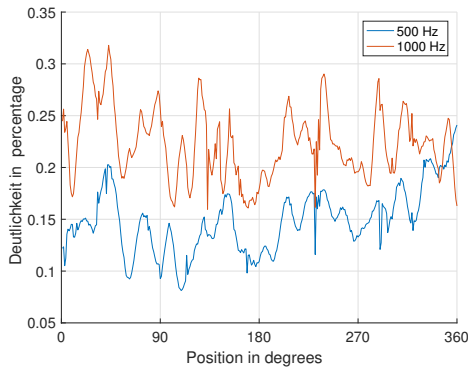
(f) Audio Lab - corner omni



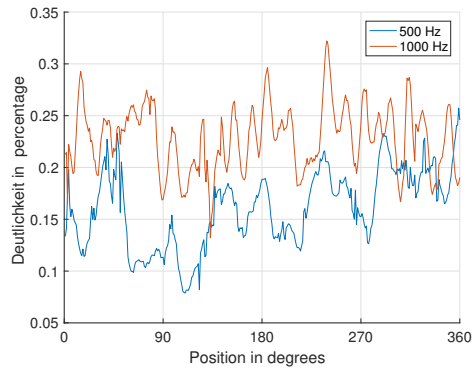
(g) Audio Lab - centre B-format



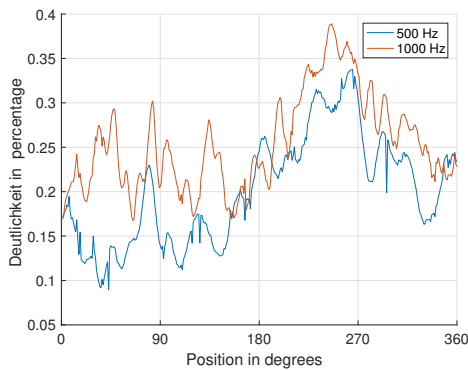
(h) Audio Lab - centre omni



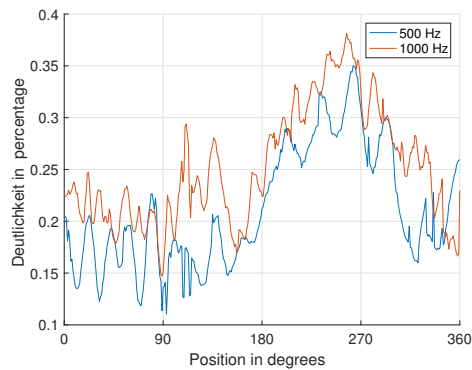
(i) Rev. chamber - corner B-format



(j) Rev. chamber - corner omni

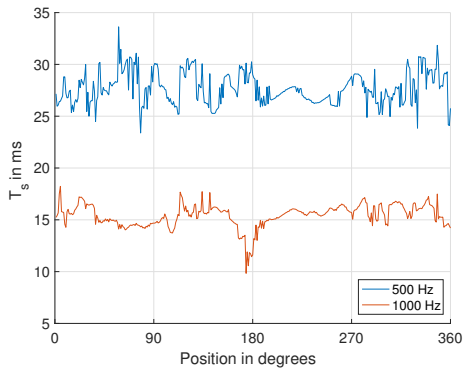


(k) Rev. chamber - centre B-format

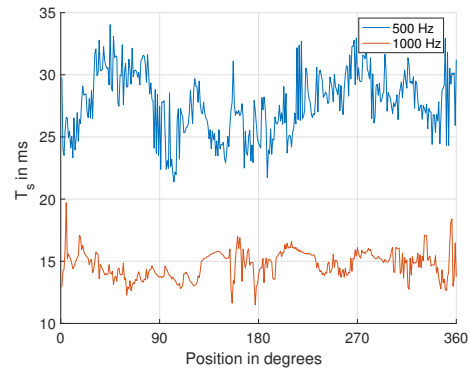


(l) Rev. chamber - centre omni

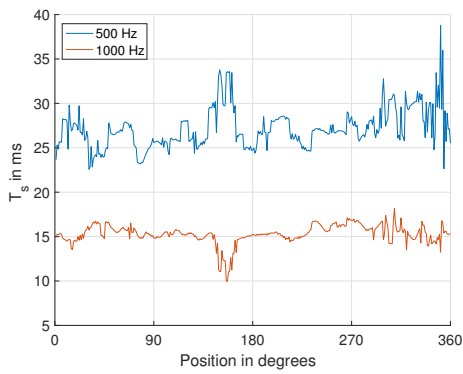
Figure 4.9: Definition D_{50}



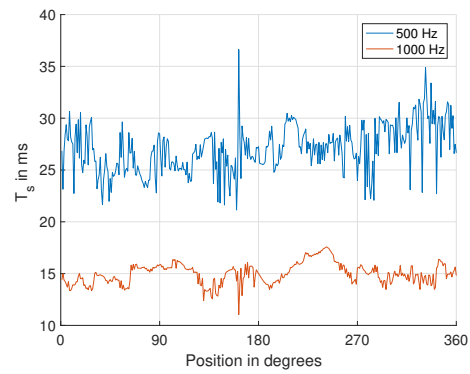
(a) Listening room corner B-format



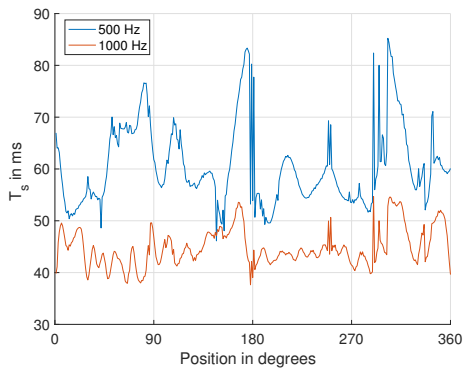
(b) Listening room - corner - omni.



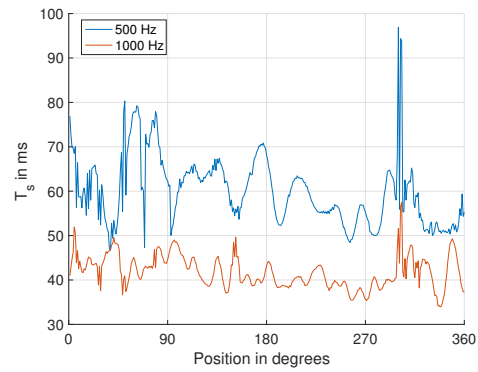
(c) Listening room centre B-format



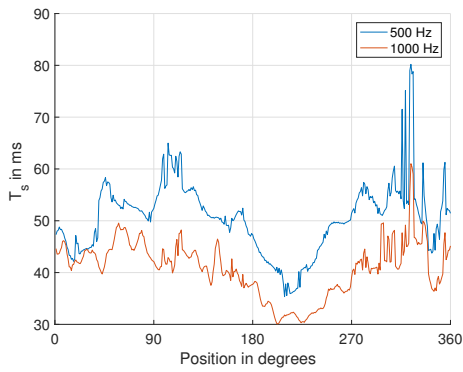
(d) Listening room centre omni.



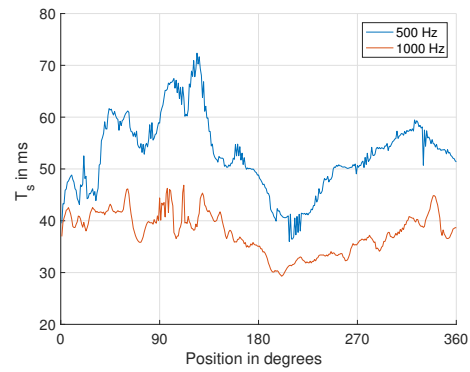
(e) Audio Lab - corner B-format



(f) Audio Lab - corner omni

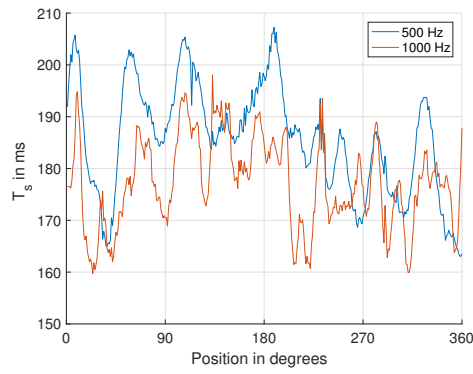


(g) Audio Lab - centre B-format

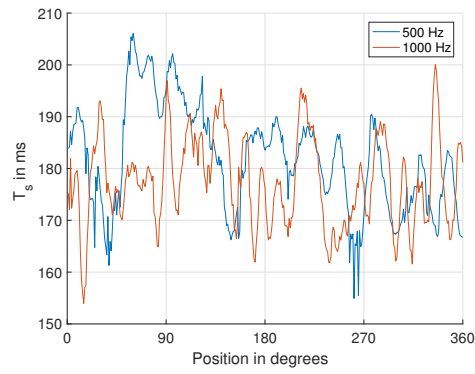


(h) Audio Lab - centre omni

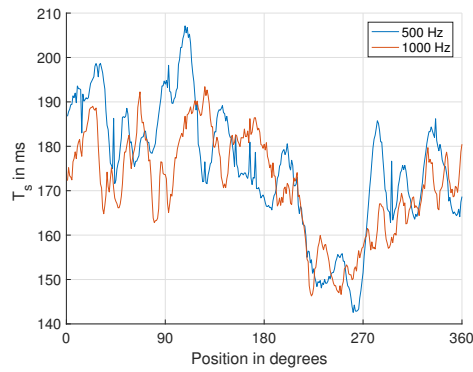
4. Results



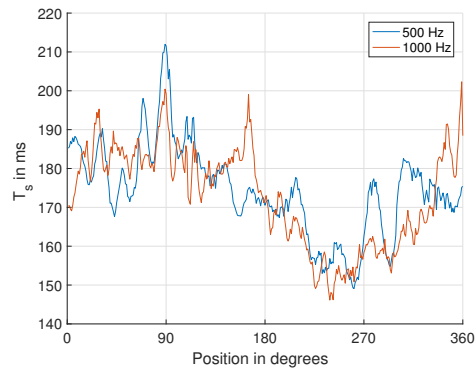
(i) Rev. chamber - corner B-format



(j) Rev. chamber - corner omni

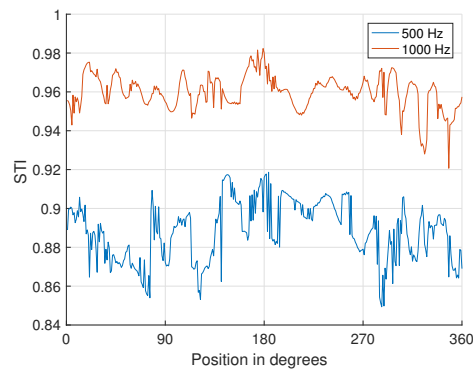


(k) Rev. chamber - centre B-format

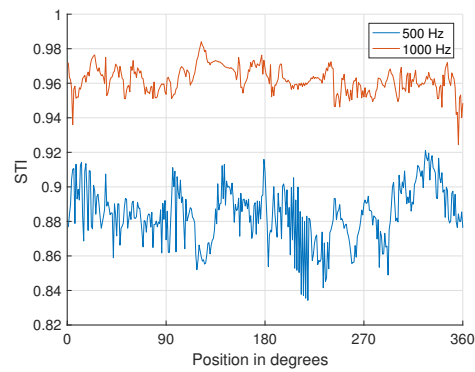


(l) Rev. chamber - centre omni

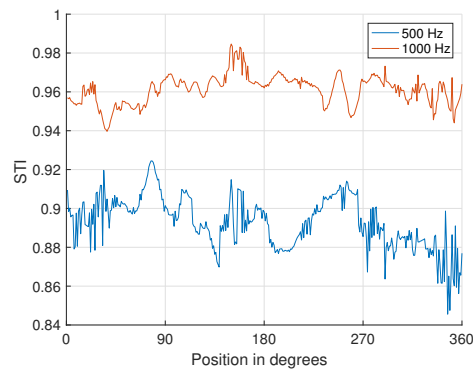
Figure 4.10: Centre Time T_s



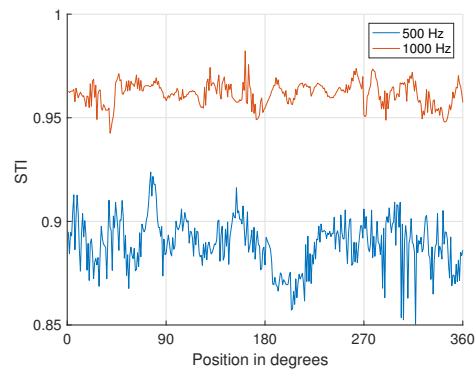
(a) Listening room corner B-format



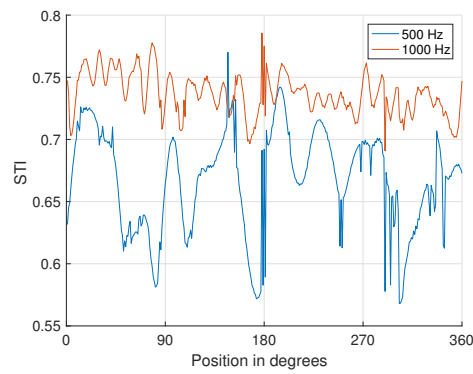
(b) Listening room - corner - omni.



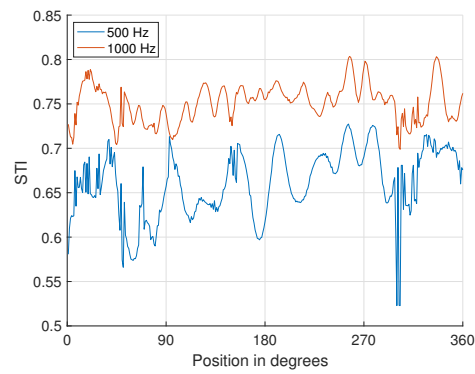
(c) Listening room centre B-format



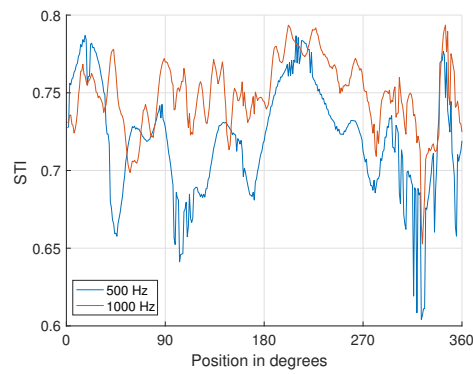
(d) Listening room centre omni.



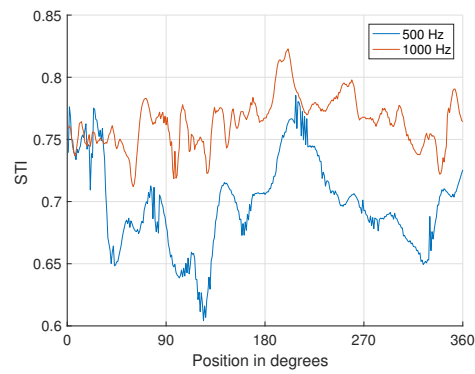
(e) Audio Lab - corner B-format



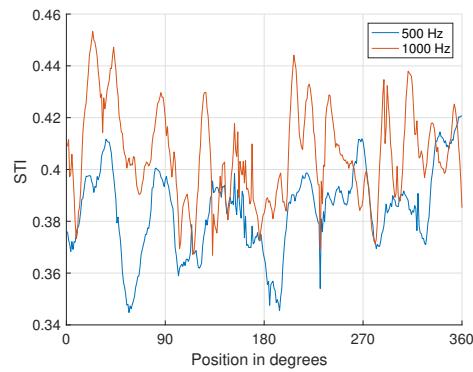
(f) Audio Lab - corner omni



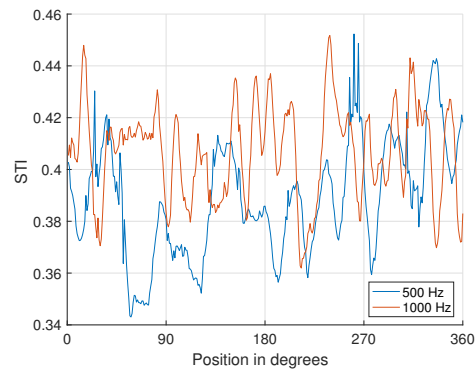
(g) Audio Lab - centre B-format



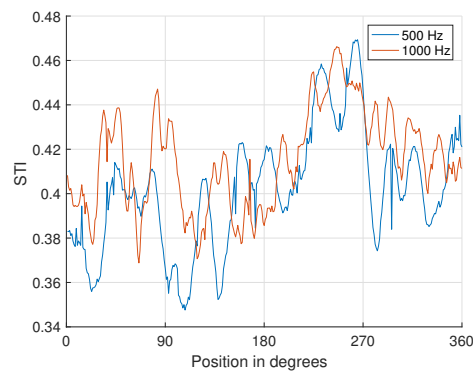
(h) Audio Lab - centre omni



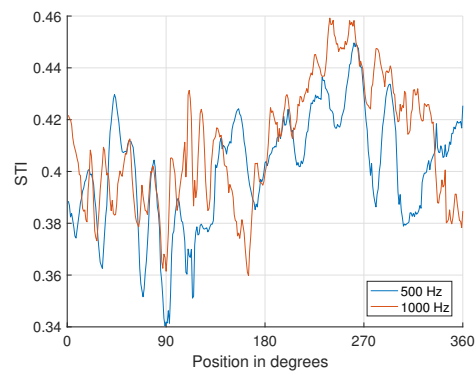
(i) Rev. chamber - corner B-format



(j) Rev. chamber - corner omni



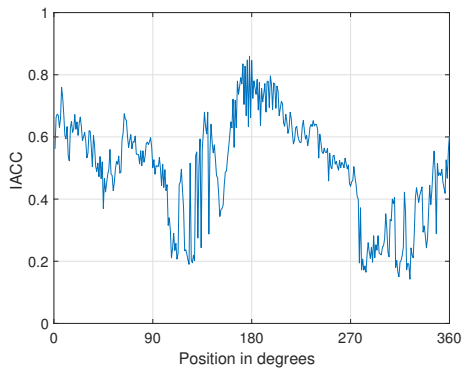
(k) Rev. chamber - centre B-format



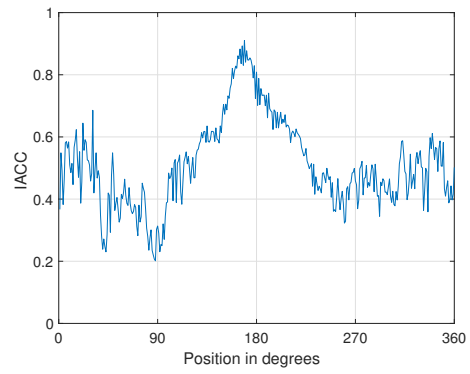
(l) Rev. chamber - centre omni

Figure 4.11: Speech transmission index STI

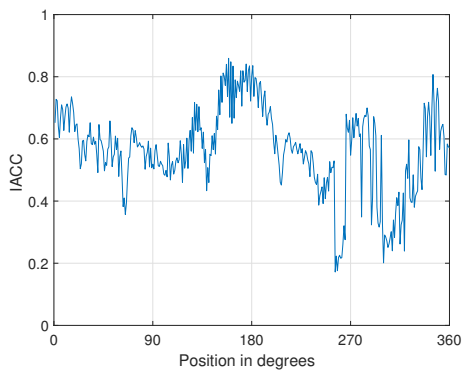
4. Results



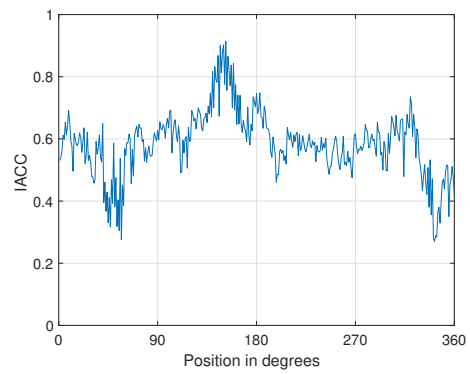
(a) Listening room corner B-format



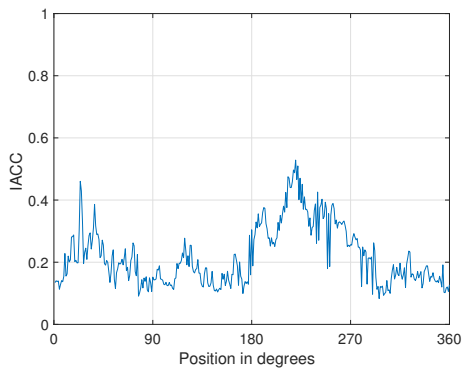
(b) Listening room - corner - omni.



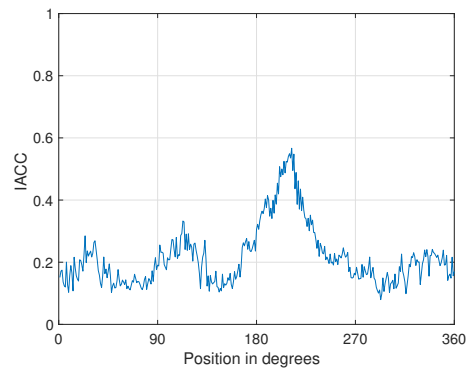
(c) Listening room centre B-format



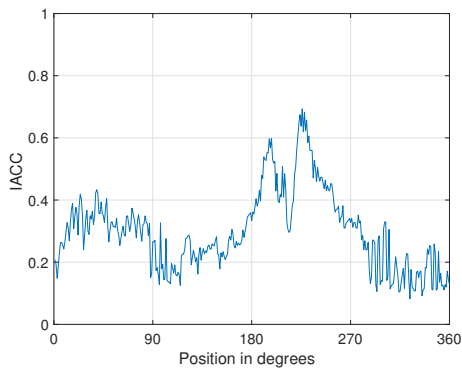
(d) Listening room centre omni.



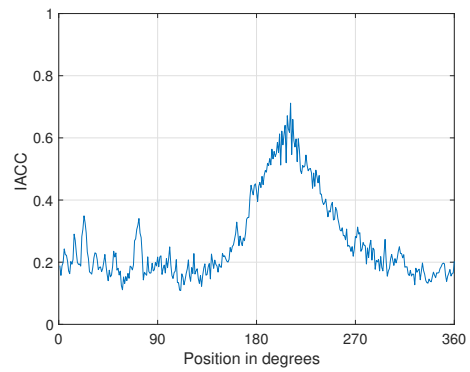
(e) Audio Lab - corner B-format



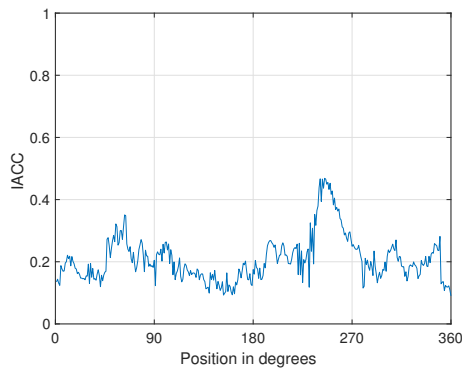
(f) Audio Lab - corner omni



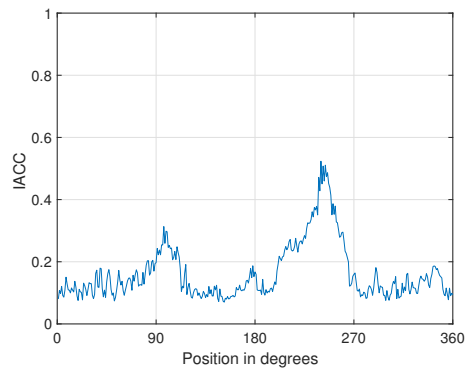
(g) Audio Lab - centre B-format



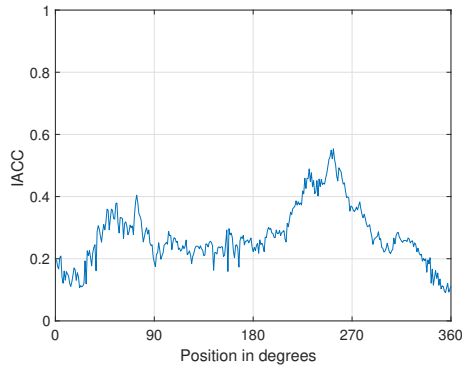
(h) Audio Lab - centre omni



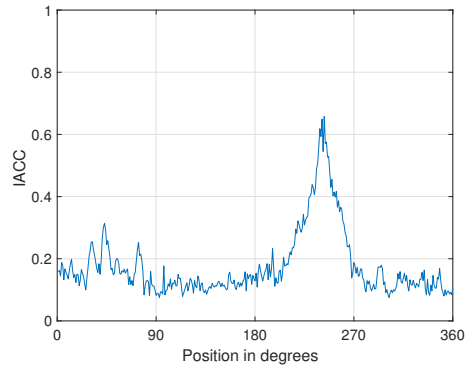
(i) Rev. chamber - corner B-format



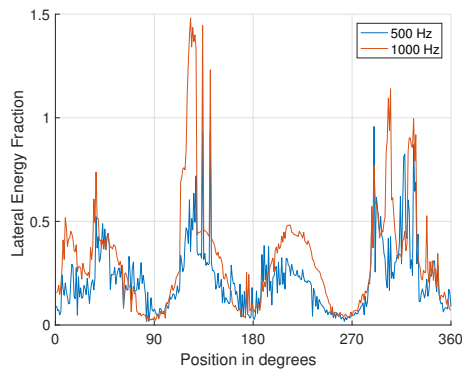
(j) Rev. chamber - corner omni



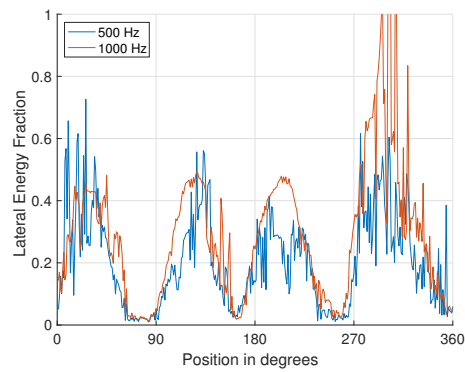
(k) Rev. chamber - centre B-format



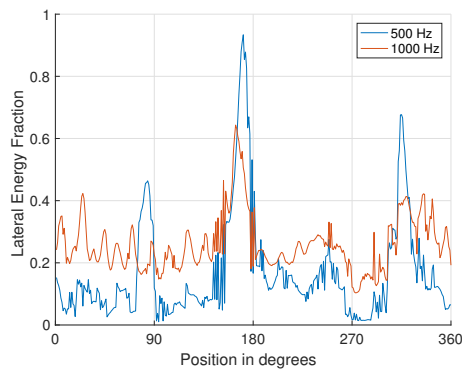
(l) Rev. chamber - centre omni

Figure 4.12: Interaural cross correlation - IACC

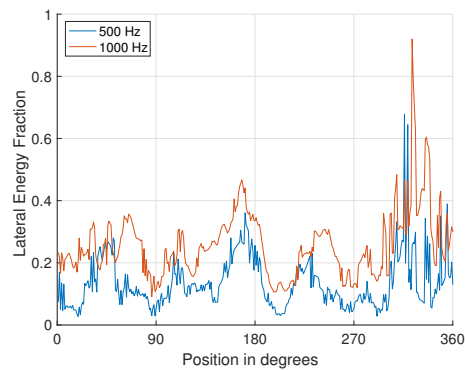
(a) Listening room - corner



(b) Listening room - centre



(c) Audio Lab - corner



(d) Audio Lab - centre

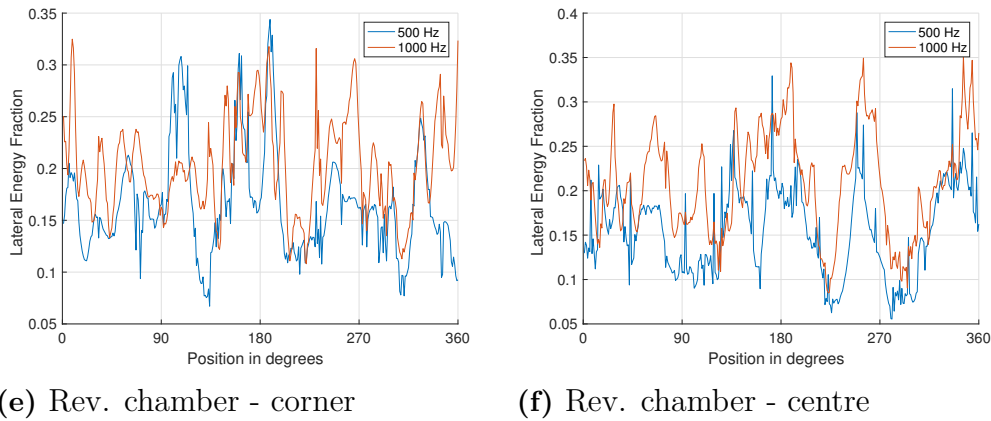


Figure 4.13: Lateral energy fraction

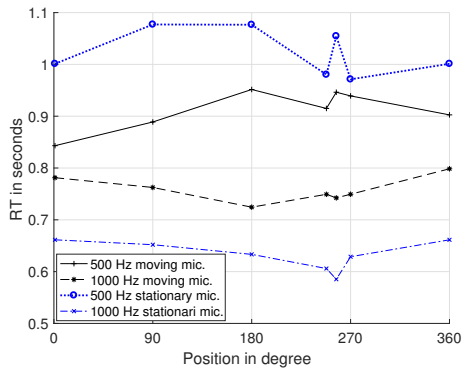
4.3 Data Comparison From Moving Microphone and Static Microphone Measurements on Equal Positions

The results of the measurements carried out with a static microphone are presented together with the results from the moving microphone system in Figure 4.14 and 4.15. Here on the left picture are presented the data measured with the B-format microphone, while the data from the omnidirectional microphone are on the right side. The points in the plots are connected with a line to permit a simpler comparison of the different trends of the resulting *curves*.

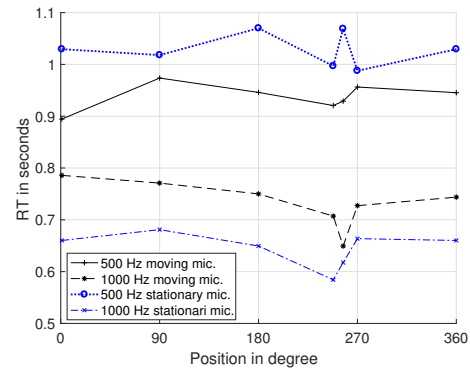
In Figure 4.14b and 4.14a is possible to recognise small but tolerable variations between the data of the two systems. The curves present also a similar trend. The other parameters in Figure 4.14 and Figure 4.15 are show similar behaviours already found in the previous sections, when the microphone is in a position in correspondence to the source. In these positions the static microphone presents even higher values than the moving microphone.

Another relevant result is shown in Figure 4.15b and 4.15a in the data at the 90° degrees position. Here the omnidirectional moving microphone system returns a high value of RT (around 3 seconds) while the static microphone computation result in a shorter RT (around 2.4 seconds). The results of the same positions obtained with the B-format microphone measurements, presents a similar value. This suggests that the isolated peaks in the T_{30} shown in Figure 4.4k and 4.4l could be a consequence of a computation error.

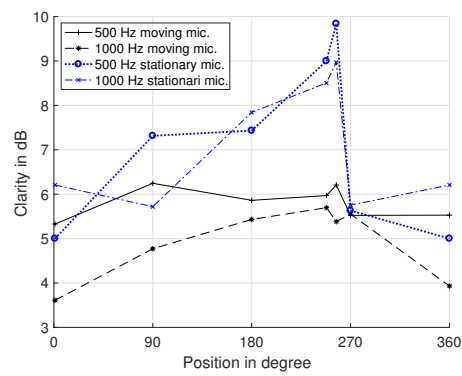
In general, the two systems in the measured positions show similar *local* data which can be considered compatible.



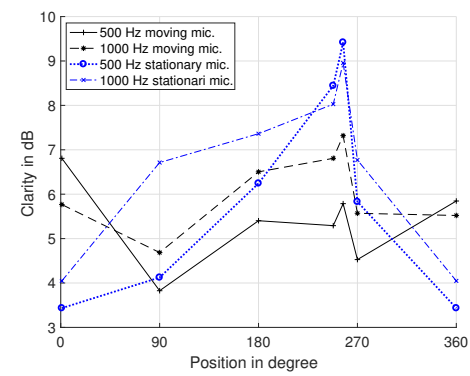
(a) T_{30} - B-format mic.



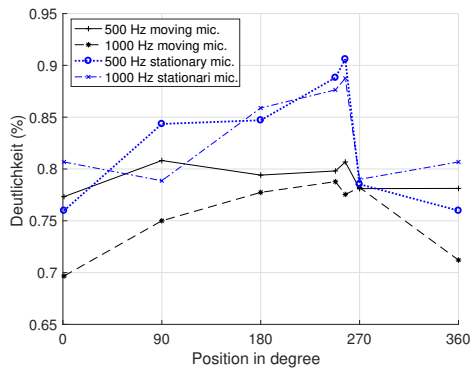
(b) T_{30} - omnidirectional mic.



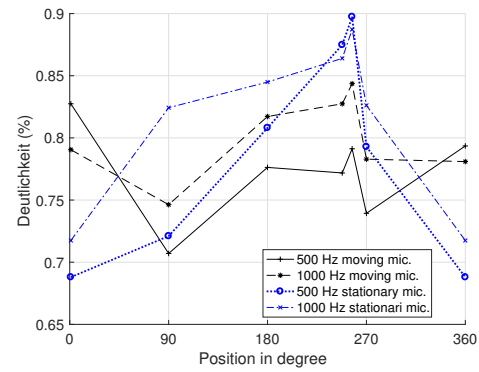
(c) C_{50} - B-format mic.



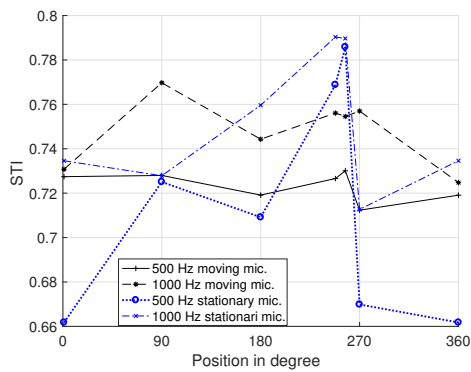
(d) C_{50} - omnidirectional mic.



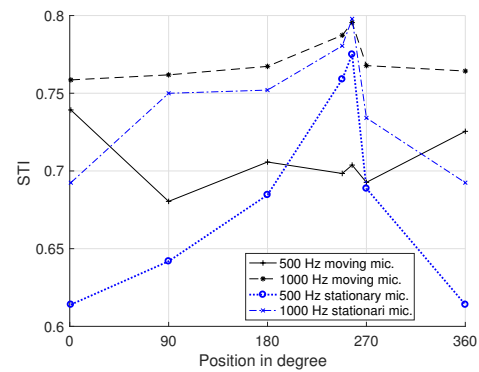
(e) D_{50} - B-format mic.



(f) D_{50} - omnidirectional mic.

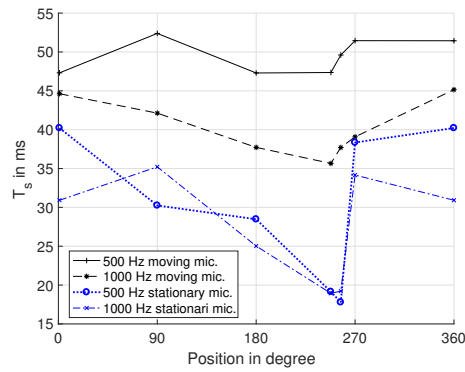


(g) STI - B-format mic.

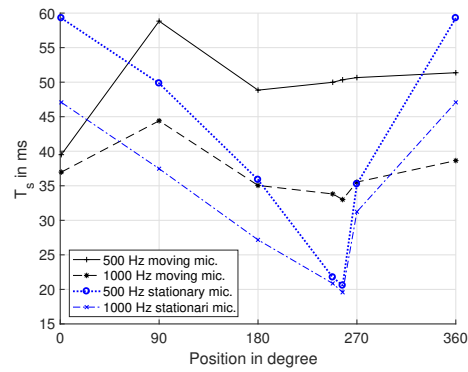


(h) STI - omnidirectional mic.

4. Results

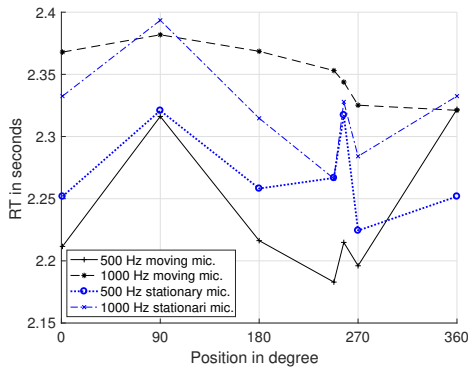


(i) T_s - B-format mic.

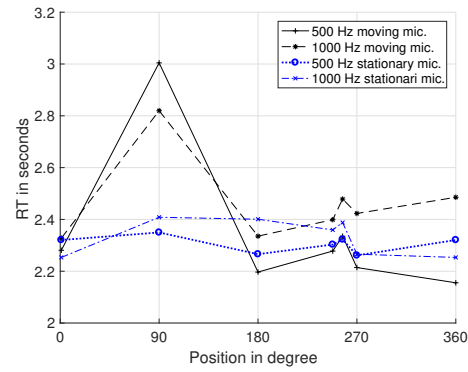


(j) T_s - omnidirectional mic.

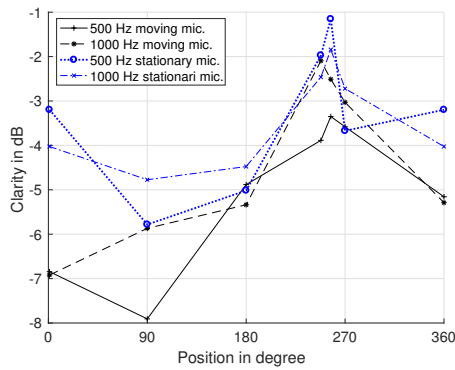
Figure 4.14: Audio Lab centre - static VS moving microphone comparison



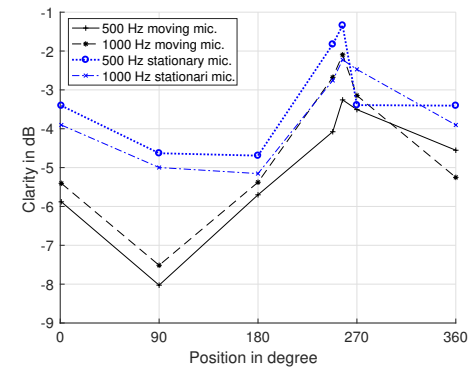
(a) T_{30} - B-format mic.



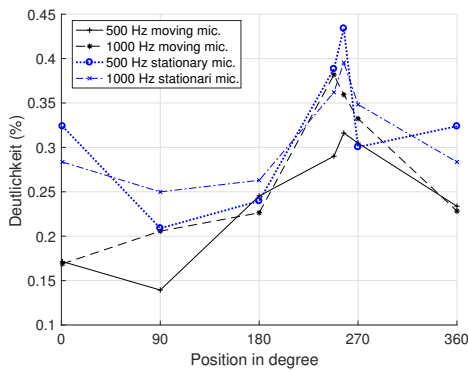
(b) T_{30} - omnidirectional mic.



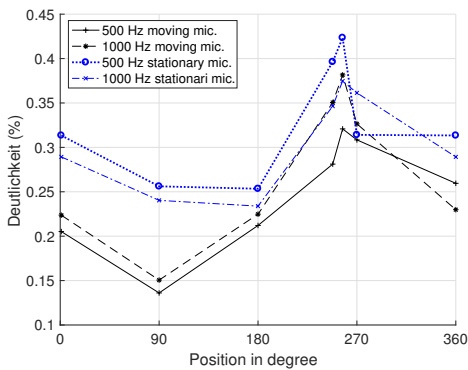
(c) C_{50} - B-format mic.



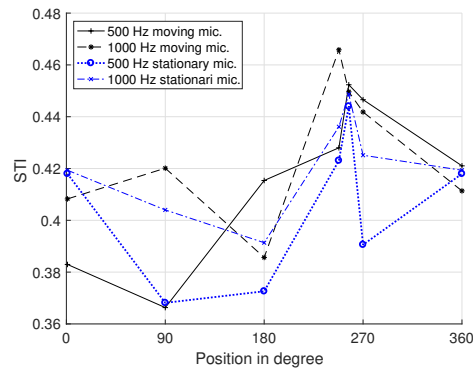
(d) C_{50} - omnidirectional mic.



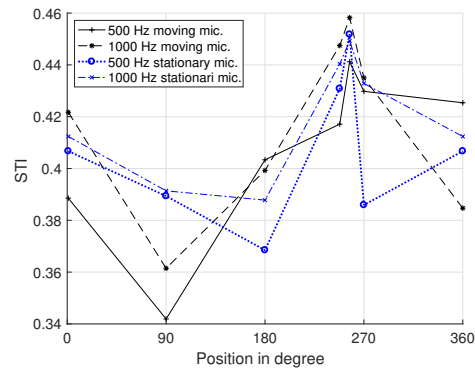
(e) D_{50} - B-format mic.



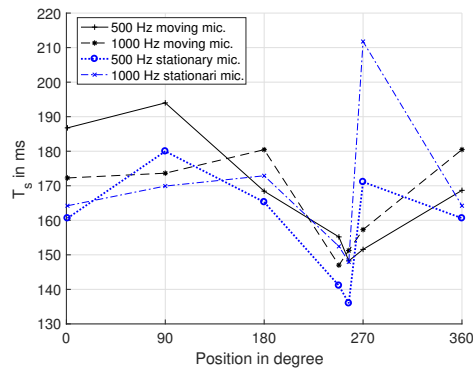
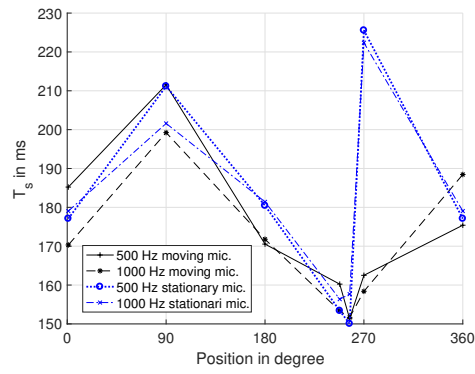
(f) D_{50} - omnidirectional mic.



(g) STI - B-format mic.



(h) STI - omnidirectional mic.

(i) T_s - B-format mic.(j) T_s - omnidirectional mic.**Figure 4.15:** Reverberation Chamber - static VS moving microphone comparison

4.4 Perceptual Evaluation

The result from the informal perceptual evaluation suggest the unreliability of the measuring system. As a consequence, a formal user study has not been considered. Since the perceptual evaluation was implemented in an informal way, without a proper questionnaire and with a limited number of testers, here only the resulting subjective impressions will be reported.

For the testers, the differences of the sound field in the three different rooms are clearly perceived. Taking into consideration of the same room, positions at significant distances are discernibly perceived. With reference to positions presented in Table 3.2, this can be observed in the IACC at 159 and 173 degrees, in the listening lab; in the reverberation time at 86 and 335 degrees in the Audio lab and in the Clarity at 173 and 309 degrees in the listening room.

While testing with sound fields at very short distance that present relevant variations in the parameters, in some case it became hard to distinguish the differences. This can be observed in the IACC at 66 and 71 degrees in the Listening room and in the reverberation time at 113 and 115 degrees in the Reverberation chamber.

This behaviour can be due to the human hearing system but it can be caused even by the measurement system. The fact that the first two reported impressions where differences at significant distances are clearly perceived, would suggest that the system is not completely reliable.

5

Conclusion and Discussion

The goal of the project was to study the local variations of room acoustics parameters. To measure the parameters a system has been used. This system acquired data continuously over a circular trajectory with a moving microphone without stops during the measurement. The collected data have been processed to obtain room impulse responses at a high spatial resolution. The system has been implemented to obtain impulses with a natural decay without noise, to simulate a binaural recording with the use of a single microphone and to be compatible with the use of standard omnidirectional microphones and a sound-field microphone.

The system has been tested in three different environments with different characteristics. If only the averaged data are considered, the system can be defined as reliable. Taking into consideration the singular results at different positions, the data show some expected behaviour over the trajectory of the measurements, but at the same time some unexpected behaviour due to probable calculation errors are observed. Finally the comparison with the perceptions obtained from the perceptual evaluation, reveal different incongruities:

while significant variations of the data are clearly perceived at large distance positions, significant variations result to be not clear or partially audible at very close positions. These episodes suggest that the singular local data show inaccuracies, turning the measurement system in a non reliable tool.

After these results, the study of the local variations cannot be properly accomplished with this methodology.

The implementation of the continuous measurement system required a much longer time than expected. The long computation time of the post-processing limited the opportunity to take different *comparative* measurements. In addition, considering the temporal limits of the project, it was not been possible to carry out a complete verification of the measurement system. Anyhow the analysis of the local data suggest a conclusion by judging the system to be not reliable, while leaving room for a possible further verification.

This does not mean that the system produces wrong data completely. It has been seen in the results that the main trend of the variations over the whole trajectory can be valid, according to the perceptual evaluation. The audio samples obtained from the convolution of the system's IRs do not present anomalous artefacts. Although the cases of local variations need more verification.

The source of the differences between the data resulting from omnidirectional microphone and B-format microphone may be found in the different filtering stages in

which the latest signals are subjected to. Other *errors* may be investigated based on the influence of the room modes in the parameters' data, an analysis which has not been considered in the project.

It must be considered that some of the measured rooms are small in size, because of which, in most of the cases the microphone positions was often close to the wall for both the cases where the VariSphear was positioned in the corner or in the centre of the room, due to the radius of the trajectory.

At the same time, as stated by De Vries in [3], more clarity about the different JND data proposed by different authors will help the interpretation of the measured data, due to the confusion in the literature.

For further, work a more solid measurement system has to be considered. If the methodology used in this project will be used, a further verification and implementation will be necessary. I will suggest to take into account the effect of the room modes and discard them from the IRs. This will discard a part of the energy but will also show how the modes affect the fluctuations of the parameters. The use of larger rooms, closer to the real cases, even with presence of furniture, can give an interesting test environment.

Bibliography

- [1] Mendel Kleiner, Jiri Tichy, *Acoustics of small room*, 2014
- [2] Johan L. Nielsen, M Miklin Halstead and A. Harold Marshall *On Spatial Variability of Room Acoustics Measures*, In Proceeding of the 15th ICA Seattle, 1998
- [3] Diemer de Vries, Edo M. Hulsebos, and Jan baan *Spatial fluctuations in measures for spaciousness*, The Journal of the Acoustical Society of America, Vol. 110 No.2, August 2001
- [4] Jens Blauert *Communication Acoustics*, Springer-Verlag Berlin Heidelberg, 2005
- [5] Thibaut Ajdler, *The plenacoustic function and its applications*, Ècole polytechnique fédérale de Lausanne, 2006
- [6] Fabrice Katzberg, Radoslaw Mazur, Marco Maass, Philipp Koch, Alfred Mertins, *Sound-field measurement with moving microphones*, The Journal of the Acoustical Society of America Vol. 114 No.5, 2017
- [7] Gerald Enzner, *3D-continuous-azimuth acquisition of head-related impulse responses using multi-channel adaptive filtering*, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, October 2009
- [8] Heinrich Kuttruff, *Room Acoustics Fifth Edition*, Spoon Press, London, 2009
- [9] International Electrotechnical Commission, *IEC-60268-16:2011 Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index*, June 2011
- [10] Behrouz Farhang-Boroujeny, *Adaptive Filters: Theory and Applications, Second Edition*, John Wiley & Sons, 2013
- [11] Jont B. Allen and David A. Berkley, *Image method for efficiently simulating small-room acoustics*, The Journal of the Acoustical Society of America, Vol. 65 No.4, April, 1979
- [12] Trevor Hastie, Robert Tibshirani, Jerome Friedman *The Elements of Statistical Second Edition Learning second edition*, Springer 2009
- [13] J.S. Bradley, R. Reich, S.G. Norcross, *A just noticeable difference in C_{50} for speech*, Elsevier - Applied Acoustics Volume 58, Issue 2 October 1999
- [14] Matthew G. Blevins, Adam T. Buck, Zhao Peng, Lily M. Wang, *Quantifying the just noticeable difference of reverberation time with band-limited noise centered around 1000 Hz using a transformed updown adaptive method*, Architectural Engineering – Faculty Publications. 63, <http://digitalcommons.unl.edu/archengfacpub/63> September 2013
- [15] Eric Benjamin, Thomas Chen *The Native B-format Microphone: Part I*, Audio Engineering Society - Convention Paper 6621 October 2005

- [16] Eric A. Lehmann and Anders M. Johansson, *Prediction of energy decay in room impulse responses simulated with an image-source model*, The Journal of the Acoustical Society of America, Vol. 124 No.1, June 2008
- [17] E. Armelloni, C. Giottoli, A. Farina, *Implementation of real-time partitioned convolution on a DSP board*, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, October, 2003
- [18] *The Auditory Modeling Toolbox*, <http://amtoolbox.sourceforge.net/models.php> [last visited January 2018]
- [19] Nicholas J. Bryan, Jonathan S. Abel, *Methods for Extending Room Impulse Responses Beyond Their Noise Floor*, Audio Engineering Society Convention Paper November 2010
- [20] I. M. Lindevald, A. H. Benade *Two-ear correlation in the statistical sound fields of rooms*, The Journal of the Acoustical Society of America Vol. 80, No.2, April 1986
- [21] Michael Brandstein, Darren Ward, *Microphone Arrays, Signal Processing Techniques and Applications*, Springer, Berlin, Heidelberg 2001
- [22] Käsbaach Johannes, May Tobias, Le Goff Nicolas, Dau Torsten *The importance of binaural cues for the perception of apparent source width at different sound pressure levels*, The journal of Annual German Congress on Acoustics (DAGA) 2014
- [23] J. Blauert, W. Lindemann, *Spatial mapping of intracranial auditory events for various degrees of interaural coherence*, Journal of the Acoustical Society of America, Vol. 79 No.3, October 1985
- [24] Benjamin Bernschütz *Microphone Arrays and Sound Field Decomposition for Dynamic Binaural Recording - Doctoral Dissertation*, Technical University of Berlin – Audio Communication Group, Technology Arts Sciences TH Köln – Institute of Communications Systems, 2016
- [25] Kilian Schulze-Forster *The B-Format – Recording, Auralization, and Absorption Measurements*, Chalmers University of Technology Gothenburg, 2018