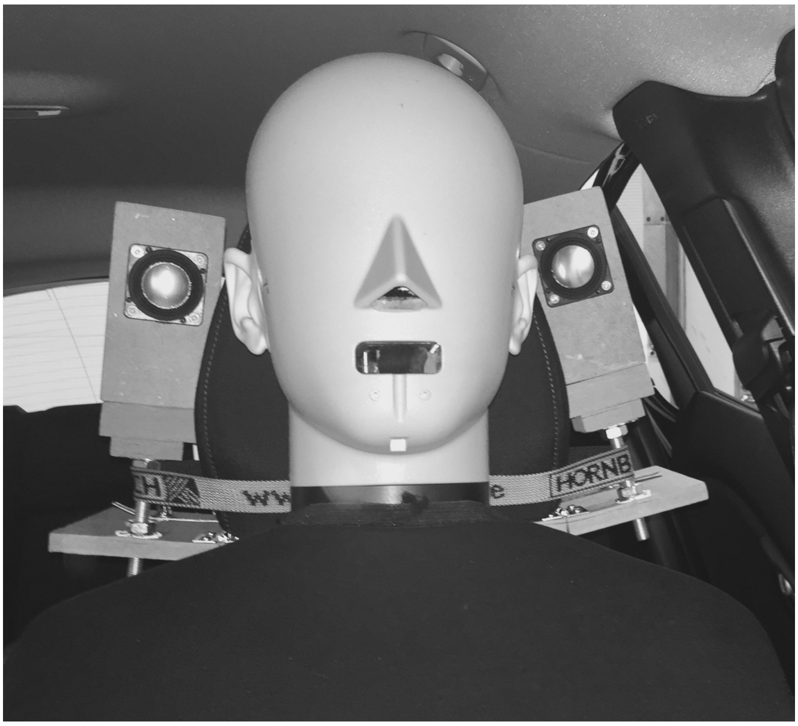
CHALMERS
UNIVERSITY OF TECHNOLOGY



Binaural sound reproduction in car compartments

A feasibility study using four channels

Master's thesis in Master Programme Sound and Vibration

JONAS KARLBERG

Division of Applied Acoustics
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2017

Master's thesis BOMX02-16-159

Binaural sound reproduction in car compartments

A feasibility study using four channels

Jonas Karlberg

Department of Civil and Environmental engineering
Division of Applied Acoustics
Chalmers University of Technology
Gothenburg, Sweden 2017

Binaural sound reproduction in car compartments
A feasibility study using four channels
Jonas Karlberg

Supervisor: Peter Mohlin, Semcon AB
Examiner: Jens Ahrens, Department of Civil and Environmental engineering, Division of Applied Acoustics

Master's Thesis BOMX02-16-159
Department of Civil and Environmental engineering
Division of Applied Acoustics
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Picture from a measurement setup using a dummy head placed inside a car compartment with custom built loudspeakers.

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2017

Binaural sound reproduction in car compartments
A feasibility study using four channels
Jonas Karlberg
Department of Civil and environmental engineering
Division of Applied Acoustics
Chalmers University of Technology

Abstract

Binaural audio material can be presented with the use of headphones or a set of loudspeakers. When using the latter, the playback system should compensate for the occurring crosstalk which is done with the use of crosstalk cancellation filters. A car compartment is considered to be acoustically challenging, for spatial reproduction over loudspeakers, due to the many reflecting surfaces. The thesis set out to investigate suitable loudspeaker positions that might be in favour for a rougher estimation of the cancellation filters, making the system more robust. Another aim was to find a method for filter design.

Firstly, a test rig with two loudspeakers was set up to try different loudspeaker positions. When the positions were set, the filters were tuned and tested with the test rig as well as inside a car compartment. Two separate systems were implemented, one behind and one in front of the listener, referred to as *back* and *front* system. When the tuning process was completed three user studies were conducted using three different loudspeaker setups: one setup using four similar loudspeakers placed in a listening room, one setup with custom built loudspeakers implemented in a car compartment and the final setup placed inside the listening room using the custom built loudspeakers from the car. The participants were asked to localise virtual source positions generated by either the back or front system.

An instrumental evaluation was performed by testing the created filters in a virtual system. Head shadowing and crosstalk performance were studied. Furthermore, a method of filter design was developed along with a measurement script for use with MATLAB or Octave.

Two of the main findings from the user study are:

- The *back* system was suitable for presenting virtual source positions behind the listener. For source positions in front of the listener most answers resulted in front/back confusions. These tendencies were shown for all tests.
- The *front* system was found to have a more uniform representation of source positions for the tests performed in the listening room. The system inside the car would need further tuning to increase system performance.

Keywords: binaural reproduction, 3D sound, filter design, car HMI, signal processing.

Acknowledgements

I would like to express my gratitude towards the entire division of Applied Acoustics; all the great people working there who are contributing to the inspiring and especially fun study environment that I experienced. I am also grateful for the employees at Alpine who participated in the listening tests as well as supplied me with *fika* every Friday.

A special thank you to Börje and Lars. Your knowledge of and the help with the measurement equipment was critical for getting started. I'm truly grateful for the time you spent helping me.

I would further like to thank Per Drougge at Alpine Electronics for all the support throughout the process of the thesis.

Thank you Peter Mohlin at Semcon for being an excellent supervisor. A true supporter who made a huge effort in guiding me along the way.

Jens Ahrens, your invaluable knowledge and positive energy helped me a lot through this project. Your passion for audio and your drive to improve the quality of the thesis was really inspiring.

Finally, I send my love to Anna, Smilla, my family and friends who have supported me all the way from when I stepped on the train to Skövde in 2010 to start this journey.

Jonas Karlberg, Gothenburg, April 2017

Contents

1 Introduction **1**
1.1 Background . . . 1
1.2 Purpose . . . 1
1.3 Limitations . . . 1
1.4 Current research . . . 2
1.4.1 3D-Sound in Car Compartments Based on Loudspeaker Reproduction Using Crosstalk Cancellation . . . 2
1.4.2 A robust algorithm for binaural audio reproduction using loudspeakers . . . 2
1.4.3 Effect of loudspeaker position on the robustness of acoustic crosstalk cancellation . . . 2

2 Theory **5**
2.1 Signal processing . . . 5
2.1.1 Deconvolution . . . 5
2.1.2 Minimum phase . . . 6
2.2 Crosstalk cancellation . . . 6
2.2.1 General CTC theory . . . 7
2.2.2 CTC using one input signal . . . 8
2.3 Estimation methods . . . 9
2.3.1 Moving average . . . 9

3 Implementation **11**
3.1 Hardware setups . . . 12
3.1.1 Test rig 0 . . . 12
3.1.2 Outcome of Test rig 0 . . . 13
3.1.3 Test rig 1 . . . 14
3.1.4 Test rig 2 . . . 14
3.1.5 Test rig 3 . . . 15
3.2 Filter design . . . 16

4 Evaluation **19**
4.1 Instrumental evaluation . . . 19
4.1.1 Back system . . . 19
4.1.1.1 Test rig 1 . . . 19
4.1.1.2 Test rig 2 . . . 20

4.1.1.3 Test rig 3 21
4.1.2 Front system 22
4.1.2.1 Test rig 1 23
4.1.2.2 Test rig 2 23
4.1.2.3 Test rig 3 25
4.2 User study 26
4.2.1 Setup 27
4.2.2 Results 27
4.2.3 Discussion 31

5 Conclusion **33**

Bibliography **35**

References **35**

A Appendix A **I**
A.1 In-vehicle measurements I
A.1.1 Placement setup 1 I
A.1.2 Placement setup 2 II
A.2 System performance and channel separation III
A.2.1 Test rig 1 III
A.2.2 Test rig 2 V
A.2.3 Test rig 3 VII
A.3 Corrected user study answers VIII
A.4 Frequency response of used loudspeakers XI
A.4.1 Loudspeakers used in user study pt.1 XII
A.4.2 Loudspeakers used in user study pt.2 and pt.3 XII

B Appendix B **XV**
B.1 Measurement scripts XV
B.1.1 ir_capture.m XV
B.1.2 create_stimuli.m XVII
B.1.3 cutIR2.m XVII
B.1.4 windctIR.m XVIII
B.2 Calculation of CTC-filters XVIII
B.2.1 ctc_calc_smooth5.m XVIII
B.2.2 limitFRF.m XXII
B.3 General XXII
B.3.1 createaudiofile.m XXII
B.3.2 setup.m XXIII
B.3.3 frfplots_general.m XXIV
B.3.4 frfplots2.m XXV
B.4 Examples XXV
B.4.1 CTC_filters.m XXV
B.4.2 CTCfilter_creation.m XXX
B.4.3 analysis.m XXXII

1
Introduction

1.1 Background

Intentional sounds have been used in cars to aid the driver since the beginning of the car industry. Audio cues such as the sound from a blinker relay switching on and off gives the driver information that the blinker is active. In recent time the technological advancements within the audio area have enabled new ways of aiding the driver with the use of 3D sound. An example of this can be to draw the driver's attention towards a certain position where a potential hazard has been identified. The car compartment is not considered to be ideal in terms of the acoustic environment needed for binaural sound reproduction, due to numerous reflective surfaces. Therefore the placement of loudspeakers are an important factor for good channel separation [1]. This thesis aims to investigate how to implement a binaural reproduction system inside a car compartment and was conducted in collaboration with Alpine Electronics, Semcon and Chalmers University of Technology.

1.2 Purpose

The following goals were defined prior to the start of the thesis:

1. To see if and how a binaural reproduction system can be implemented in a car
2. To investigate suitable loudspeaker positions for placement inside a car
3. To find a suitable method for filter design for use in a binaural playback system

Binaural reproduction systems have been implemented in cars in previous research, but the method for the actual filter design and derivation of the filter equations are not fully disclosed in the relevant literature. Therefore, the thesis also sets out to present the reader with a method of designing their own binaural playback system.

1.3 Limitations

Some authors emphasise the need for individual head related transfer functions, *HRTFs*, when tuning the system for optimal playback [2] [1]. Individual HRTFs were not implemented within the scope of this thesis. To further optimise the playback system, the cancellation filters could be set to adapt for head movement. However, the investigation, implementation and evaluation of the system presented in this thesis will only be for one listening position. A suggestion of how to deal with this problem is to use spatial averaging when creating the crosstalk cancellation filters [3].

1.4 Current research

This section includes short summarises of the most relevant papers within the topic.

1.4.1 3D-Sound in Car Compartments Based on Loudspeaker Reproduction Using Crosstalk Cancellation

Authors: Andre Lundkvist, Arne Nykänen and Roger Johnsson.

The paper describes testing of binaural loudspeaker configurations inside a car compartment. Tests were performed in a car compartment and in an anechoic environment. It was found that the crosstalk cancellation filters were most effective in the range of 2500-4000 Hz. Good channel separation was achieved with loudspeakers placed close behind the head of the listener and this was probably due to the headphone-like setup. For further development the authors suggest standard stereo techniques in addition to the binaural playback system in order to achieve better localisation. It is also suggested that the system should be calibrated for each individual and that the played sounds should be recorded inside the car to further improve localisation. Another suggestion is to try loudspeaker positions behind the head at farther distances [1].

1.4.2 A robust algorithm for binaural audio reproduction using loudspeakers

Authors: W. Jie et.al.

The paper investigates the robustness of the channel separation when the listening position changes i.e. when the listener's head moves. The authors propose a multi-position weighted method i.e. spatial averaging for dealing with the mentioned problem. They present a method for finding the optimal filter with respect to head movement. However, further studies are needed for choosing the proper positions to create such filters [3].

1.4.3 Effect of loudspeaker position on the robustness of acoustic crosstalk cancellation

Authors: D. Ward and G. Elko

The letter describes a way to find the optimal positions for placement of loudspeakers for use in a binaural playback system. They started with two closely placed loudspeakers, referred to as the *stereo-dipole*. They found that the spacing between the loudspeakers should vary with frequency in an ideal case. Wider loudspeaker spacing is desirable for lower frequencies and for frequencies above 4 kHz a spacing of $\pm 5^\circ$ is desirable. The practical implementation of this would be $\pm 20^\circ$ for low frequencies and $\pm 5^\circ$ for high frequencies [4].

2

Theory

The concept of using other sources to control a sound field, globally or locally, originated from an idea presented in 1936. Paul Lueg presented his idea of Active Noise Control, *ANC*, in the patent *"Process of silencing sound oscillations"* [5]. While the idea is quite simple, the implementation of such systems has proven to be difficult without robust and efficient hardware and algorithms. The use of digital signal processors, *DSPs*, and signal processing algorithms have progressed over the years and at present time the technology has found its way to consumer electronics.

The following sections briefly describe the theory used throughout the thesis. For further reading, more in depth theory can be found in the cited paper and the references therein.

2.1 Signal processing

A linear time invariant system, LTI-system, is a mathematical model describing the features of a system. One property of such system is that the order of the filtering processes does not affect the resulting signal [6]. The following subsections give a description of the necessary steps which were used in the filter design.

2.1.1 Deconvolution

If an unknown LTI system is described by the impulse response $h(t)$ its properties can be found by using an input signal $x(t)$ and measuring the output $y(t)$, see figure 2.1. The output of the system is the convolution of the input signal and the system impulse response as seen in equation 2.1 [7].

$$x * h = y \tag{2.1}$$

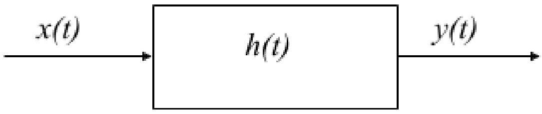


Figure 2.1: A signal $x(t)$ sent into system $h(t)$ with its output $y(t)$.

To find the transfer function $H(\omega)$ for the system, one can use the expression seen in equation 2.2 by computing the Fourier transform for x and y [7].

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} \tag{2.2}$$

2.1.2 Minimum phase

A discrete system is considered to be stable if the poles and zeros, in a zero-pole representation, are located inside the unit circle. The system $H(z)$ can be described by the fraction of two polynomials $A(z)$ and $B(z)$ as seen in equation 2.3 [7].

$$H(z) = \frac{B(z)}{A(z)} \tag{2.3}$$

The zeros of each polynomial define the zeros in the numerator and poles in the denominator. An important property of a stable filter, with its poles and zeros inside the unit circle, is that the inverse of the system, is stable as well [7].

If a system with its frequency response $H(\omega)$ has poles outside the unit circle, the minimum phase of the system $H_{mp}(\omega)$ have the same frequency response, but with polynomials with zeros and poles inside the unit circle. The condition is seen in expression 2.4 [7][8].

$$|H(\omega)| = |H_{mp}(\omega)| \tag{2.4}$$

2.2 Crosstalk cancellation

When listening to binaurally prepared audio, the playback is usually performed with the use of headphones, since no further processing of the signals are needed. Using a set of headphones delivers the signal to its intended receiver, i.e. the left and right ear of the listener, without any substantial amount of crosstalk between the ears. When using a set of loudspeakers one needs to cancel out the right source signal that is received at the left ear and vice versa. This is done by designing crosstalk cancellation filters. The following sections explain the basic principles of crosstalk cancellation, referred to as *CTC*. An explanation of the following expressions is seen in the list below, all in the frequency domain.

X_i input signal with index i, i.e. left or right
Y_i signal at the receiver with index i, i.e. left or right
H_{ij} transfer function of the propagation path from source i to receiver j

2.2.1 General CTC theory

The received signals at each ear can be described by the matrices as seen in equation 2.5 [3].

$$\begin{bmatrix} Y_L \\ Y_R \end{bmatrix} = \begin{bmatrix} H_{LL} & H_{RL} \\ H_{LR} & H_{RR} \end{bmatrix} \cdot \begin{bmatrix} X_L \\ X_R \end{bmatrix} \tag{2.5}$$

In an ideal case, the resulting signal at the listeners ears is equal to the input signal i.e. as in equation 2.6.

$$\begin{bmatrix} Y_L \\ Y_R \end{bmatrix} = \begin{bmatrix} X_L \\ X_R \end{bmatrix} \tag{2.6}$$

In equation 2.5 it is clear that the matrix H, containing the transfer paths, needs to be compensated for with an additional filter matrix. The design criterion for such filter matrix C, can be described by equation 2.7, where the multiplication of the transfer paths and the desired filter results in an identity matrix[1].

$$\begin{bmatrix} H_{LL} & H_{RL} \\ H_{LR} & H_{RR} \end{bmatrix} \cdot \begin{bmatrix} C_{LL} & C_{RL} \\ C_{LR} & C_{RR} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{2.7}$$

The needed filter matrix C is then H^{-1} and is given by equation 2.8 [1].

$$C = \frac{1}{H_{RR}H_{LL} - H_{RL}H_{LR}} \begin{bmatrix} H_{RR} & -H_{RL} \\ -H_{LR} & H_{LL} \end{bmatrix} \tag{2.8}$$

In figure 2.2 the signal path of a two channel binaural playback system is described. The system can be described in matrix form as shown in equation 2.9.

$$\begin{bmatrix} Y_L \\ Y_R \end{bmatrix} = \begin{bmatrix} H_{LL} & H_{RL} \\ H_{LR} & H_{RR} \end{bmatrix} \cdot \begin{bmatrix} C_{LL} & C_{RL} \\ C_{LR} & C_{RR} \end{bmatrix} \cdot \begin{bmatrix} X_L \\ X_R \end{bmatrix} \tag{2.9}$$

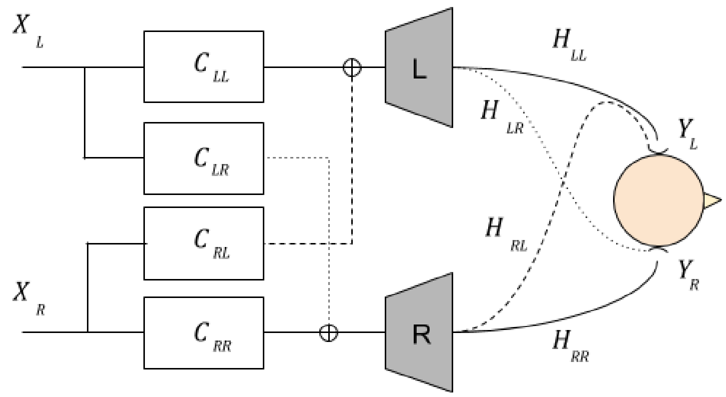


Figure 2.2: Overview of a two channel reproduction system with input signals X_i, CTC-filters C_{ij}, loudspeakers L & R, propagation paths H_{ij} and received signals Y_i.

When matrix C equals H^{-1} the condition in 2.6 is fulfilled.

2.2.2 CTC using one input signal

In this section a derivation of the expressions for C is presented. To simplify the resulting expressions the derivation is performed for one input channel. The desired filters can be derived by looking at the total signal received at each ear. The input signal is described by X_L. In figure 2.3 an overview of the signal path is seen.

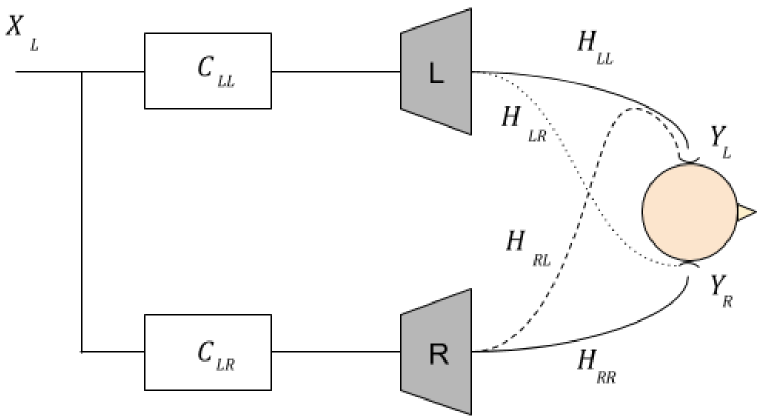


Figure 2.3: One channel signal path.

The resulting signal at the ears are seen in equation 2.10.

$$Y_L = X_L C_{LL} H_{LL} + X_L C_{LR} H_{RL} \qquad = X_L(C_{LL} H_{LL} + C_{LR} H_{RL}) \tag{2.10a}$$
$$Y_R = X_L C_{LL} H_{LR} + X_L C_{LR} H_{RR} \qquad = X_L(C_{LL} H_{LR} + C_{LR} H_{RR}) \tag{2.10b}$$

Which gives the conditions seen in equation 2.11.

$$C_{LL}H_{LL} + C_{LR}H_{RL} = 1 \tag{2.11a}$$
$$C_{LL}H_{LR} + C_{LR}H_{RR} = 0 \tag{2.11b}$$

The equation system is solved as seen in the following steps which give us the filters C_{LL} and C_{LR}.

Solving 2.11a for C_{LL} gives

$$C_{LL} = \frac{1 - C_{LR}H_{RL}}{H_{LL}}. \tag{2.12a}$$

Using 2.12a in 2.11b gives

$$C_{LR}H_{RR} + \frac{1 - C_{LR}H_{RL}}{H_{LL}}H_{LR} = 0 \Leftrightarrow C_{LR}\frac{(H_{LL}H_{RR} - H_{RL}H_{LR})}{H_{LL}} = -\frac{H_{LR}}{H_{LL}}$$

$$C_{LR} = \frac{-H_{LR}}{H_{LL}H_{RR} - H_{RL}H_{LR}}. \tag{2.12b}$$

Using 2.12b in 2.12a gives

$$C_{LL} = \frac{H_{RR}}{H_{LL}H_{RR} - H_{RL}H_{LR}}. \tag{2.12c}$$

When adding more input signals the desired filters are obtained using the described method.

2.3 Estimation methods

When the propagation paths are measured the CTC-filters can be calculated as shown in previous sections. However, when using the strictly theoretical model the filters will only be valid for the position in which the measurements were performed. This means that the system will be very fragile against head movements. The measurements will likely have narrow dips in the frequency spectrum, and when calculating H^{-1} these dips will be narrow peaks instead which will cause the system to fail [3].

To tackle these challenges the filters need to be estimated in some way which in the end will be a trade off between robustness and performance.

2.3.1 Moving average

A method to smooth the frequency response is to use a moving average filter. The operation is described as

$$y(i) = \frac{1}{N}\sum_{j=0}^{N-1} x(i+j) \tag{2.13}$$

where x is the input signal, y the output of the operation and N the number of points used in the average [9]. The implementation of this can easily be achieved by using the built in function *smooth* in MATLAB curve fitting toolbox [10].

3
Implementation

The development and implementation were the result of continuous testing, tuning and evaluating system performance by investigating different parameter settings. The implementation chapter describes the core steps performed. More information related to the implementation can be found in *Appendix A*.

All measurements performed during the thesis used the following equipment, if not stated otherwise.

- **B&K type 4100** - Head and torso simulator, *HATS*
- **B&K type 2804** - Battery driven power supply
- **M-audio ProFire 610** - Sound card
- **Behringer reference amplifier model A500** - Power amplifier
- **Custom built power amplifier** - Multichannel power amplifier
- **Custom built amplifier** - Amplifier with gain 0 to +40 dB, steps of 20 dB
- **Custom built loudspeakers** - a total of 8 loudspeakers were built
- **Unit T model UT71B** - Multimeter for checking output voltage of the power amplifiers

The general connection setup for the measurements is presented in figure 3.1.

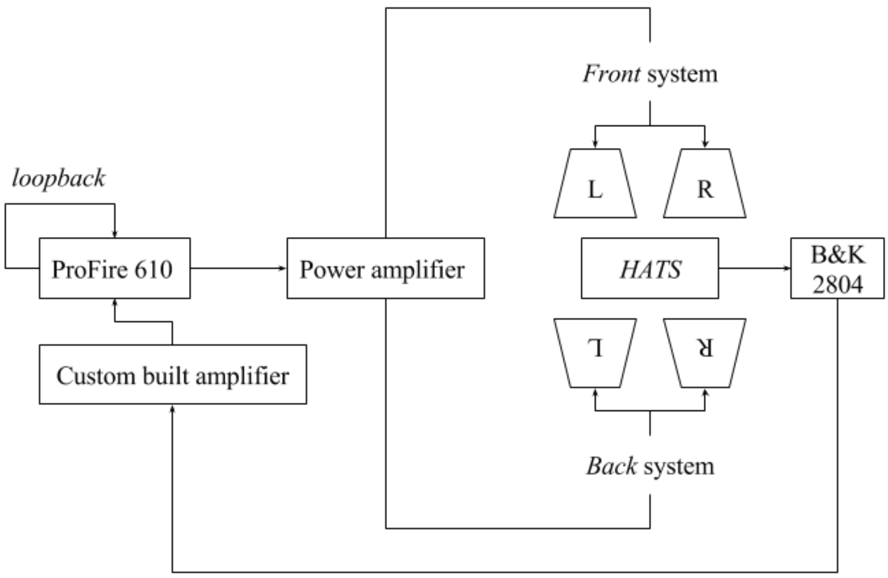


Figure 3.1: Block diagram of connections for the measurement setups.

A script in Matlab was written for measuring the impulse responses for the different setups. The script is found in *Appendix B*. To validate the measurements performed with the script additional measurements were performed with the use of *Room Equalization Wizard*, known as *REW*. The measurements made using the script were considered to be reliable since both the script and REW gave the same output.

For the frequency responses of the loudspeakers captured using REW, see *Appendix A*.

3.1 Hardware setups

A total of four test rigs were implemented with different purposes. A brief description of each system is found in the following sections.

3.1.1 Test rig 0

As a starting point for finding a suitable way of designing filters, a pre-study of suitable positions was conducted. The main goal was as follows:

- to get a functioning setup for binaural reproduction in an reverberant environment
- to try different speaker positions; discover pitfalls when choosing positions inside the car
- to conduct tests when evolving the filter design
- to evaluate the performance of the filters
- to measure the effects of head shadowing
- to optimise the Matlab/Octave code

Test rig 0 was set up in a combined listening and conference room at Alpine's office. The setup consisted of the dummy head, a chair, two loudspeakers mounted on stands and the Behringer power amplifier. The room was left furnished and the front and side walls had some additional absorptive material. In figure 3.2 a picture shows one of the tested positions.

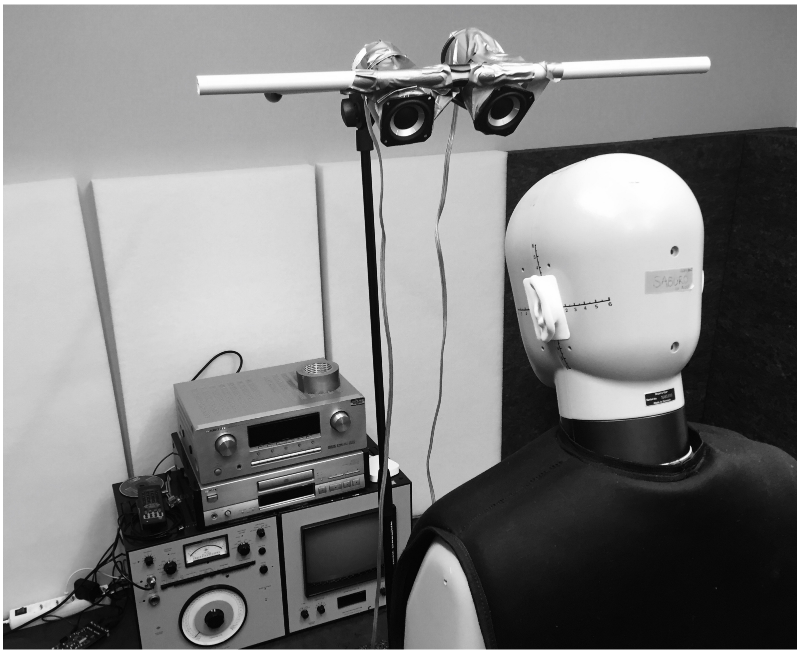


Figure 3.2: Picture from one of the loudspeaker positions used in *Test rig 0*.

3.1.2 Outcome of Test rig 0

Several different loudspeaker positions were investigated for both the back and front systems. The positions were chosen to mimic positions that could be implemented inside the car compartment as well as to investigate effects of different loudspeaker distances. Each loudspeaker configuration was measured with an ideal listening position being symmetrically placed between the loudspeakers as well as rotating the head to the left. The head was turned to the left in approximate 1 cm steps. Markers were placed on the neck of the *HATS* for reference when rotating the head. When evaluating the measured responses, a Dirac impulse was used as left input and sent into a virtual system[1]. The right input signal was set to zero. The input signals were then fed through the measured responses and calculated filters. The resulting reduction was calculated using the measured responses, primarily looking at the effects of head shadowing. The head shadowing curves were obtained by calculating the difference between the direct path, i.e. signal to the ear, and the crosstalk path. For the loudspeaker configurations that seemed to have good channel separation, binaural soundfiles were played through the system and the robustness of the system was tested.

It was found that positions to the side and behind the head had good separation due to the strong effects of head shadowing. However, the crosstalk cancellation works well where it is needed i.e. below 1 kHz and in the range of 3-4 kHz as stated in [1].

For the positions placed in front of the listener, good crosstalk cancellation was achieved, but with low robustness.

[1]The *virtual system* was implemented as a script in MATLAB.

Based on these trials it was decided that speakers would be mounted in the ceiling with a speaker separation distance of 150 mm and side by side at the headrest with a driver to ear distance of 140 mm.

3.1.3 Test rig 1

Test rig 1 was created to evaluate the filter designs and the influence of the room. Four custom built loudspeakers were used in the trials. The setup was used in *User study pt.1* which is further described in section 4.2. In figure 3.3 the setup is presented.

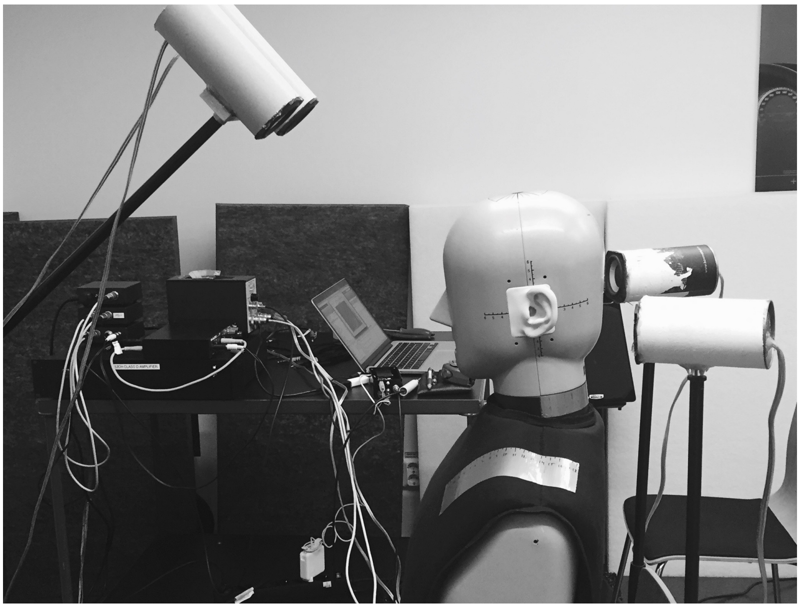


Figure 3.3: Picture from setup *Test rig 1*.

3.1.4 Test rig 2

Test rig 2 was implemented inside a Volvo V40 with two custom built loudspeakers for use in the back system and two drivers mounted inside the ceiling with the car ceiling acting as a baffle. The setup was used in *User study pt.2* as well for further testing and tuning of the filters.

Robustness was checked by measuring several different head positions. Firstly, the head was rotated left and right in 1 cm steps following the reference points placed at the neck of the *HATS*. Then additional positions were measured with the whole *HATS* slanted to the left and right as well as towards the steering wheel. In figure 3.4 the rotation and slanting are illustrated.

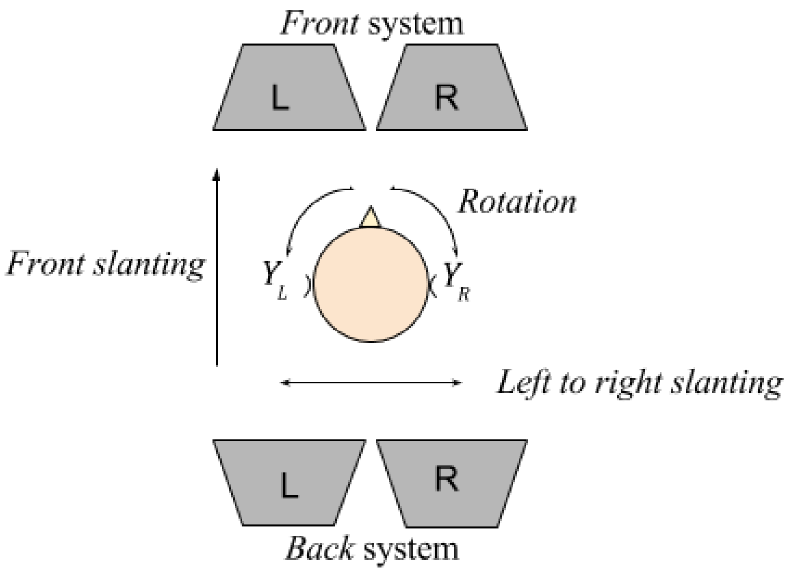


Figure 3.4: Robustness check, rotation and slanting of the dummy head.

A detailed description of the measurements performed for *Test rig 2* is found in *Appendix A*. In figure 3.5 the setup is presented.

Figure 3.5: Picture from setup *Test rig 2*. The right front driver is seen mounted in the ceiling in the top right corner of the picture.

3.1.5 Test rig 3

The final setup, *Test rig 3*, consisted of the four loudspeakers from *Test rig 2* with the front system drivers mounted in a custom built speaker box tuned to match the frequency response when mounted in the car ceiling. The loudspeakers were placed in the same room as *Test rig 1*. The placement setup was similar to *Test rig 1* but with other loudspeakers. This configuration was used in *User study pt.3*. A picture of the setup is seen in figure 3.6.

Figure 3.6: Picture from setup *Test rig 3*. The picture was taken before beginning *User study pt.3*.

3.2 Filter design

As previously mentioned, the filter design was the outcome of continuous testing and tuning of filter parameters. The testing was performed with the use of *Sound scape renderer*. The SSR software ran with two simulatinous instances, the *Generic renderer* for live convolution of the calculated filters and the *Binaural renderer* performning convolution of hrirs[2] for binaural playback. In figure 3.7 a simple sketch of the signal flow using SSR is seen.

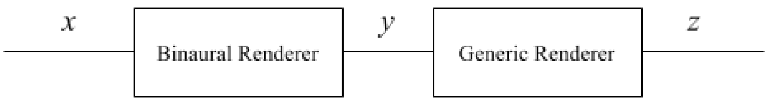


Figure 3.7: Signal flow of SSR instances, with x being the input audio signal, y the binaural audio output signal and z the output of convolution with crosstalk filters.

With the calculation of crosstalk cancellation filters in matrix C in expression 2.8 a way to tackle the many reflections from inside the car compartment is to cut the measured impulse responses and thereby obtaining a semi anechoic environment i.e. altering the room or transmission channel. Then the calculated filters are compensating for a simplified version of the room which serves as a sort of smoothing. By changing the length of the impulse responses, the frequency resolution is altered.

[2] Head related impulse responses are convolved with the input signal thus placing the virtual sound source in a sound stage.

To estimate the colouration effects of the playback system, informal listening tests were carried out.
In figure 3.8 the resulting filter design is shown. The MATLAB code for the filter creation is found in *Appendix B* and the general signal flow is described below.

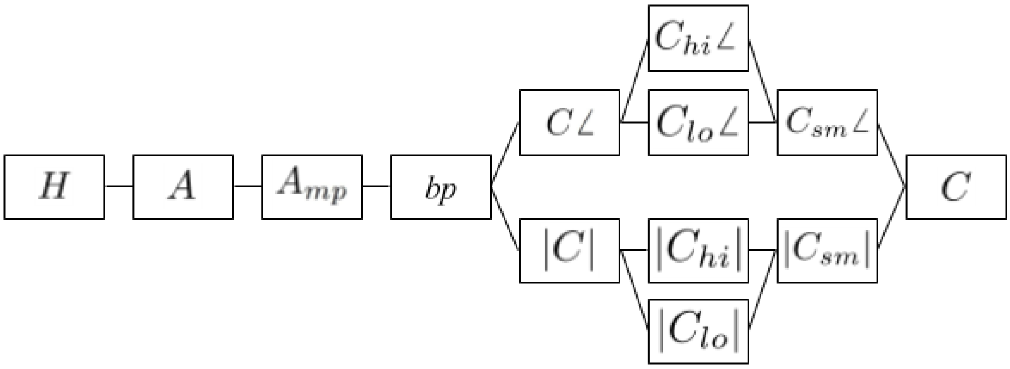


Figure 3.8: Flow chart of the final filter design steps. H is the measured frequency responses; A is the calculation of the denominator; A_{mp} is the minimum phase response of A; *bp* is a bandpass filter; $C\angle$ contains the phase information from the calculated filters; $C_{hi}\angle$ and $C_{lo}\angle$ denote the splitting of the phase vector for different smoothing settings; $|C|$ is the magnitude response of the calculated filters; $|C_{hi}|$ and $|C_{lo}|$ denote splitting of the magnitude response; $|C_{sm}|$ is the smoothed magnitude response; $C_{sm}\angle$ is the smoothed phase response and finally C is the resulting filter.

The measured FRFs are loaded into the filter creation script and there the denominator of expression 2.8, A, is calculated. Since A will be inverted, the minimum phase response of A is calculated and based on this stable response the filter matrix C is obtained. The filters are band pass filtered and windowed in time domain before separating the magnitude and phase. The band pass filter is a combination of two Butterworth filters[3] and the window operation is a Hanning window with different slopes for fade in and fade out. With the response separated, both phase and magnitude are split into two frequency regions to perform smoothing using a moving average with different values of N. This is done in order to be able to smooth an upper and lower frequency range with N_i. When the moving average filter operation is applied the whole frequency range is smoothed again to even out any abrupt changes in the frequency response where the splitting is merged. Finally, the phase is reconstructed and thereby the filter design is complete.

When creating the filters a maximum dynamic range of 20 dB was set as a design criterion. The trade-off between robustness, colouration and channel separation was tuned by listening to the system and trying different smoothing settings.

[3]The two filters creating the band pass filter: A sixth order high pass filter with f_c at 180 Hz and a twenty fourth order low pass filter with f_c at 12000 Hz.

4

Evaluation

This chapter includes the results and analysis of the user study. The systems' crosstalk performance and other insights are also presented. For measured loudspeaker responses see *Appendix A*.

4.1 Instrumental evaluation

When evaluating the performance of the different systems the channel separation between the ears was studied. For both systems this includes separation in total, due to the created filters and the effects of head shadowing. The frequency response for the illuminated ear in perfect conditions should be a flat line at 0 dB. The values have been rounded in the text corresponding to the general trend of the plots. All plots have been smoothed with a moving average of 500 points for visibility. Plots of the raw data is presented in *Appendix A*.

4.1.1 Back system

The channel separation for the back system is generally good for all setups. The advantageous placement of the speakers improves the separation where, in some cases, the filters contribute in a destructive way, in terms of channel separation.

4.1.1.1 Test rig 1

For *Test rig 1* the total channel separation was at best -35 dB and in the frequency range from 100 to 1000 Hz the separation was -7 to -18 dB. Figures 4.1 and 4.2 show the crosstalk performance and channel separation. The filters contribute with -3 to -8 dB in different frequency spans and at worst the negative contribution ranges from an average of +4 dB up to a peak value of +9 dB at 6731 Hz. The frequency response for the illuminated ear has some colouration, but within a reasonable range of ±3 dB throughout the response up to 3000 Hz. From 4200 Hz and up a gain peak of 5 dB and a 4 dB dip can be seen.

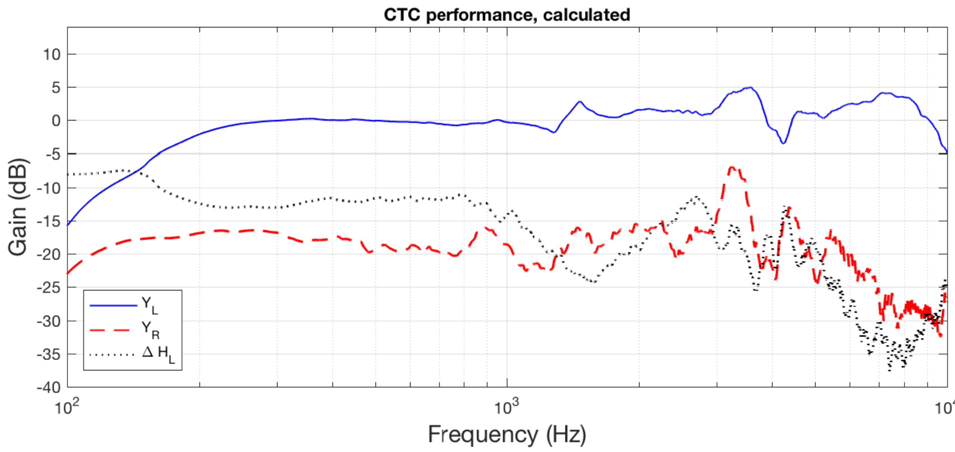


Figure 4.1: Frequency response with crosstalk performance for *back* system, *Test rig 1*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

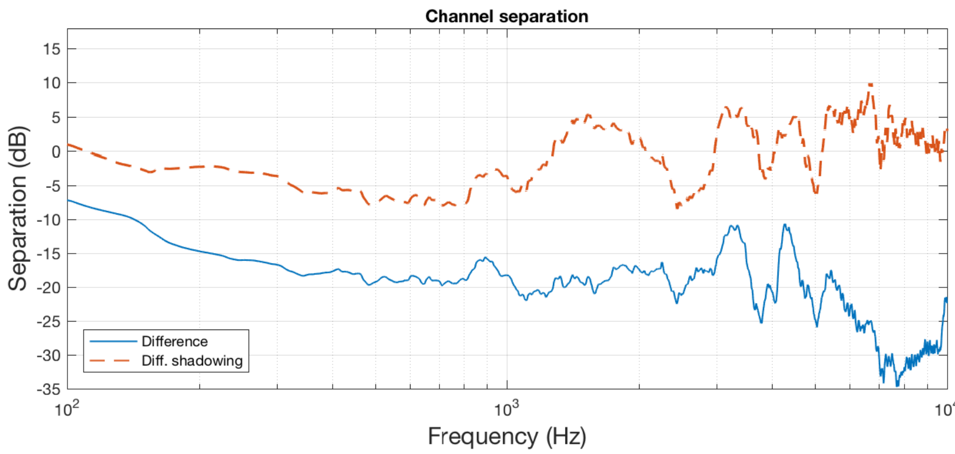


Figure 4.2: Channel separation for *back* system, *Test rig 1*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line, *Diff. shadowing*, shows the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

4.1.1.2 Test rig 2

The total channel separation was at best -29 dB at 5862 Hz. As seen in figures 4.3 and 4.4 the filters contribute the most in the frequency ranges 100 to 300 Hz and 450 to 1300 Hz ranging from -10 dB to -2 dB and -3 dB to -14 dB respectively. From 2300 Hz and up the filters decrease the channel separation up to +14 dB at 8411 Hz. Exceptions in the upper frequency range are found from 3500 Hz to 4020 Hz and 5700 Hz to 6700 Hz where the filters increase channel separation up to -7 dB and -3 dB respectively. For the frequency response of the illuminated ear a flat response is obtained from 205 Hz to 1300 Hz. From 1570 Hz and up a repeatable pattern is seen with increasing peaks and dips up to a maximum of +12 dB and -10 dB. At 1570 Hz the wavelength corresponds to 0.22 m. The peaks and dips could then be due to a strong reflection omitting from the left speaker and side wall of the listener. A reflection in time domain causes a repeatable pattern in frequency domain, causing reoccurring dips in the frequency response[1].

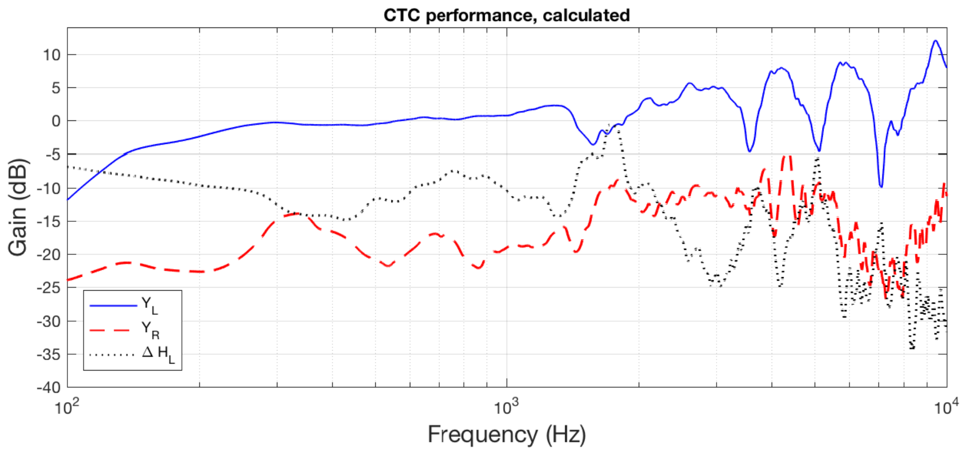


Figure 4.3: Frequency response with crosstalk performance for *back* system, *Test rig 2*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

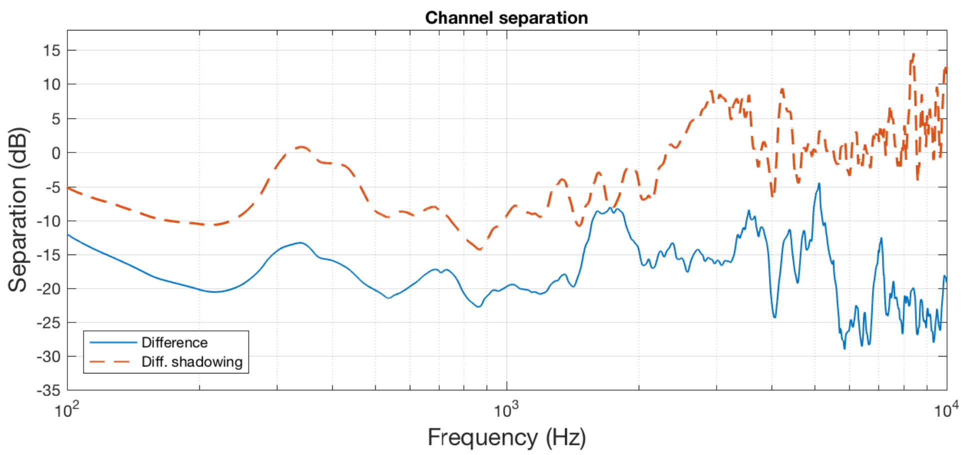


Figure 4.4: Channel separation for *back* system, *Test rig 2*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line, *Diff. shadowing*, shows the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

4.1.1.3 Test rig 3

In figure 4.5 and 4.6 the performance and channel separation for *Test rig 3* are shown. Here, a total channel separation of -28 dB was achieved at 8240 Hz. The filters contribute up to -15 dB in the frequency range from 100 to 1000 Hz and up

[1]An impulse with a reflection occurring at time $\Delta t = t_2 - t_1$ (s) is seen as repeated dips in the frequency response occurring at every $\Delta f = f_2 - f_1$ (Hz).

to -7 dB from 1700 to 2850 Hz. As with previous setups the negative contribution occurs mainly at higher frequencies with a max value of +15 dB. Up to 2000 Hz the spectrum is within ±1 dB. The strongest peak and dip occur at 3600 and 4300 Hz respectively with gain changes of +5 dB and -11 dB.

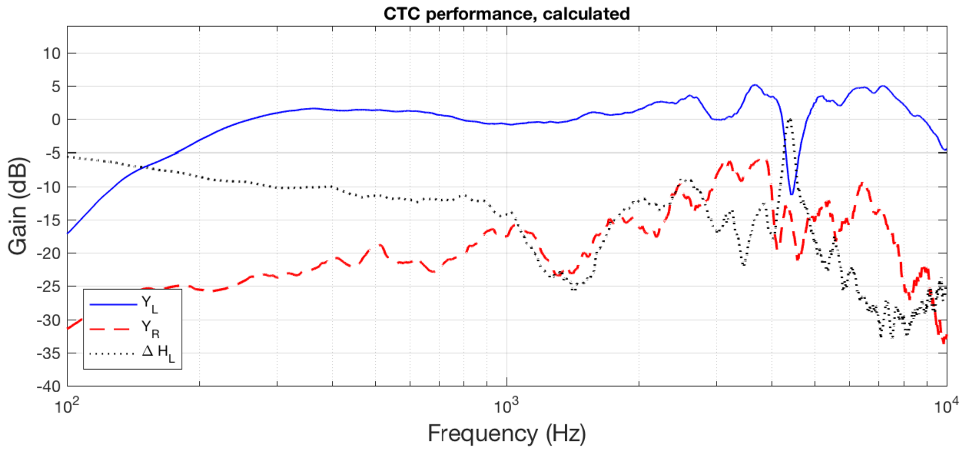


Figure 4.5: Frequency response with crosstalk performance for *back* system, *Test rig 3*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

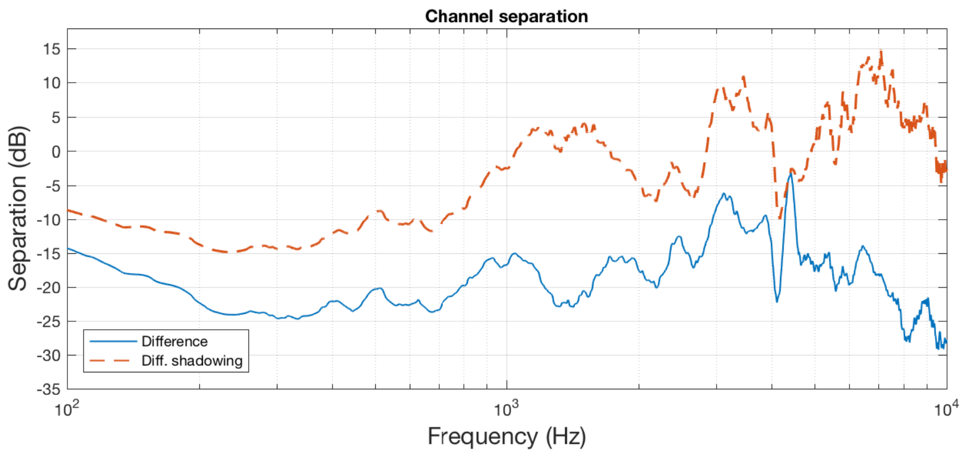


Figure 4.6: Channel separation for *back* system, *Test rig 3*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line, *Diff. shadowing*, shows the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

4.1.2 Front system

Due to the placement of the front system loudspeakers, the contribution from head shadowing in terms of performance is small. As seen in the following plots, the channel separation is mainly due to the filters.

4.1.2.1 Test rig 1

Figure 4.7 shows the crosstalk performance and figure 4.8 the channel separation. The total channel separation was at best -26 dB at 7492 Hz. The filters contribute with -20 dB at the most at around 3000 Hz. For frequencies above 8500 Hz, when head shadowing starts to improve, the filters have a negative contribution up to +8 dB. The frequency response for the illuminated ear has strong attenuation up to 300 Hz and gain up to +10 dB from 2000 to 5000 Hz.

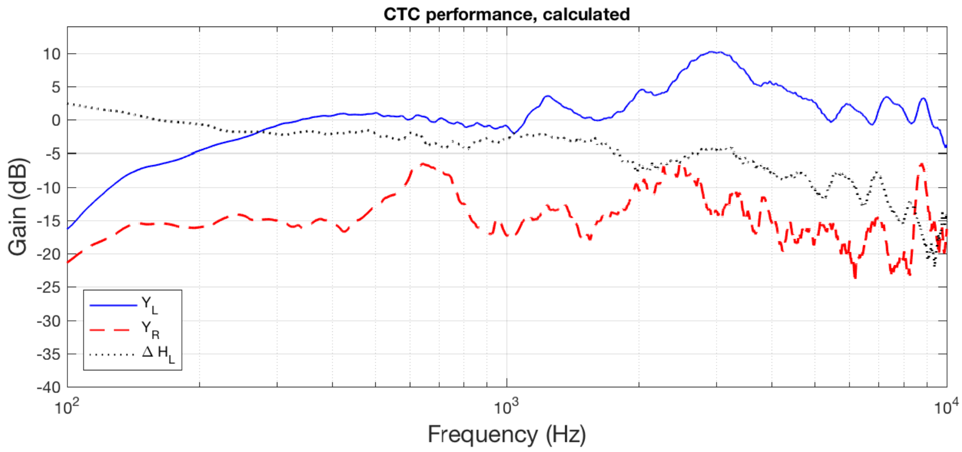


Figure 4.7: Frequency response with crosstalk performance for *front* system, *Test rig 1*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

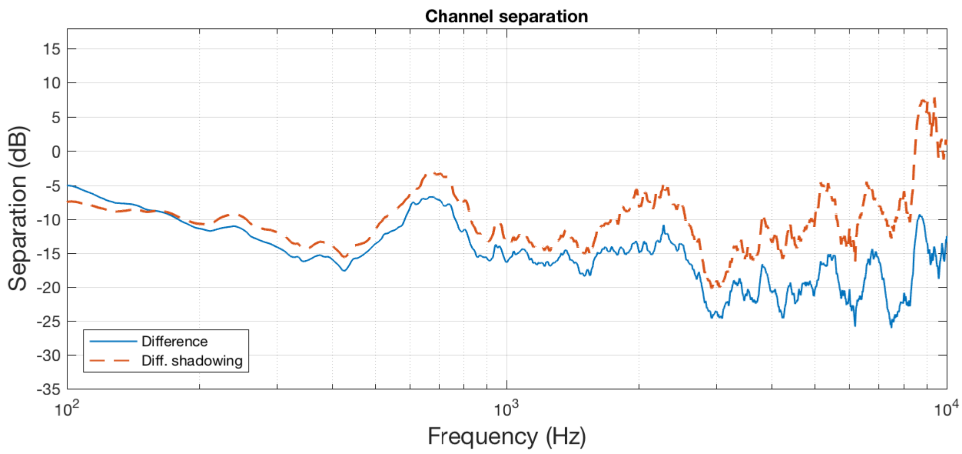


Figure 4.8: Channel separation for *front* system, *Test rig 1*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line, *Diff. shadowing*, shows the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

4.1.2.2 Test rig 2

In figure 4.9 and 4.10 the performance and channel separation for *Test rig 2* are shown. In this setup the frequency response as well as channel separation has more narrow peaks and dips. The total channel separation was at best -17 dB. Up to 900 Hz the filters contribute with -12 dB and in the frequency range from 1300 to 2000 Hz the attenuation is -10 dB. The filters' negative contribution were at worst +16 dB at 8060 Hz. The frequency response has a total dynamic of ±15 dB with the highest peak up to +11 dB and dip -5 dB. Since the filters are compensating for the frequency response of the loudspeakers, the simplification of the created filters in addition to the many reflecting surfaces could be an explanation for the resulting response.

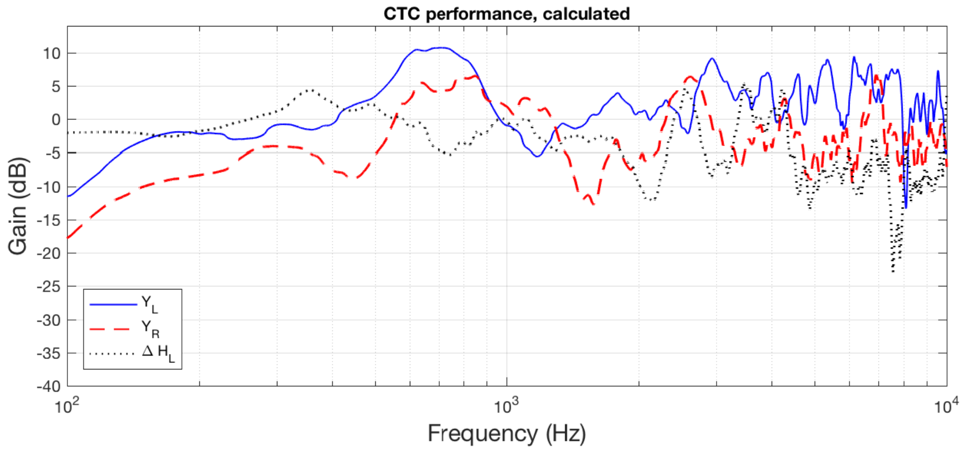


Figure 4.9: Frequency response with crosstalk performance for *front* system, *Test rig 2*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

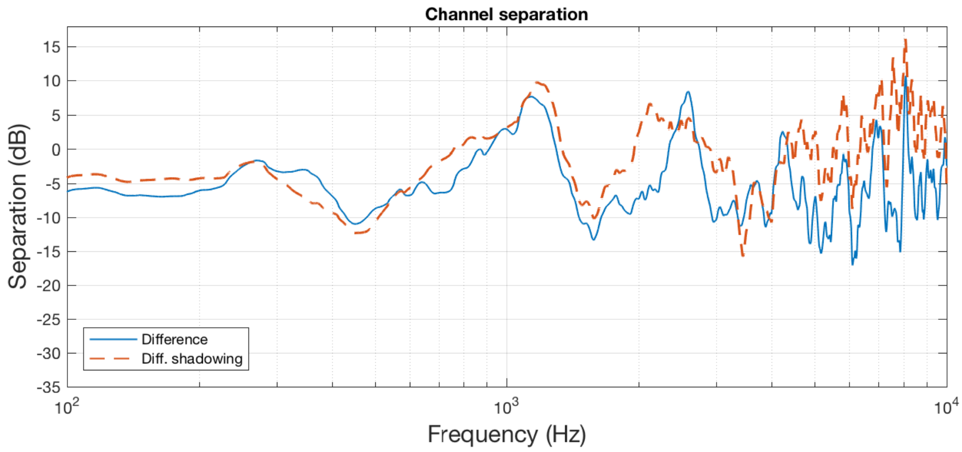


Figure 4.10: Channel separation for *front* system, *Test rig 2*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line, *Diff. shadowing*, shows the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

4.1.2.3 Test rig 3

The total channel separation was at best -25 dB as seen in figure 4.12. The filters start contributing at 170 Hz with the highest contribution of -17 dB at 1900 Hz. The filters continue to increase the channel separation up to around 7450 Hz where a negative contribution of +10 dB occurs. In figure 4.11 the frequency response of the illuminated ear stays within ±2 dB from 150 to 2300 Hz. Around 2300 Hz, peaks of +6 dB are present. Looking at the different frequency responses for the *front* system, *Test rig 3* has the flattest resulting response.

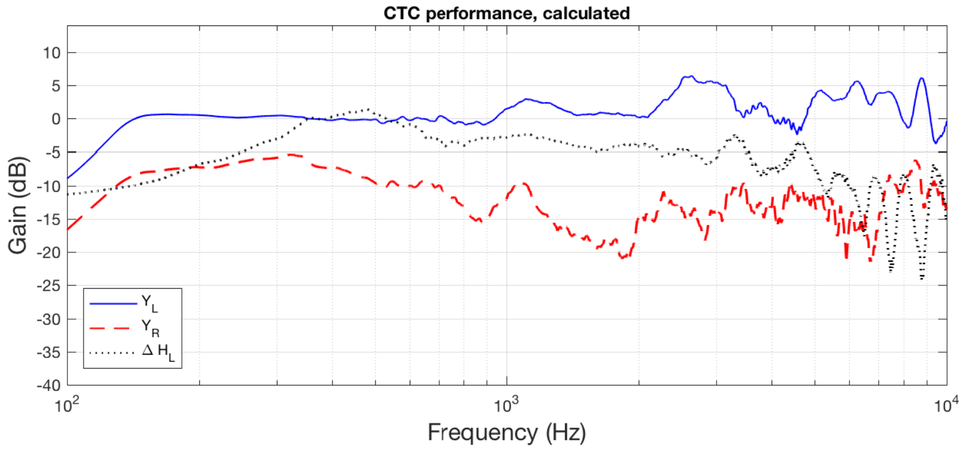


Figure 4.11: Frequency response with crosstalk performance for *front* system, *Test rig 3*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

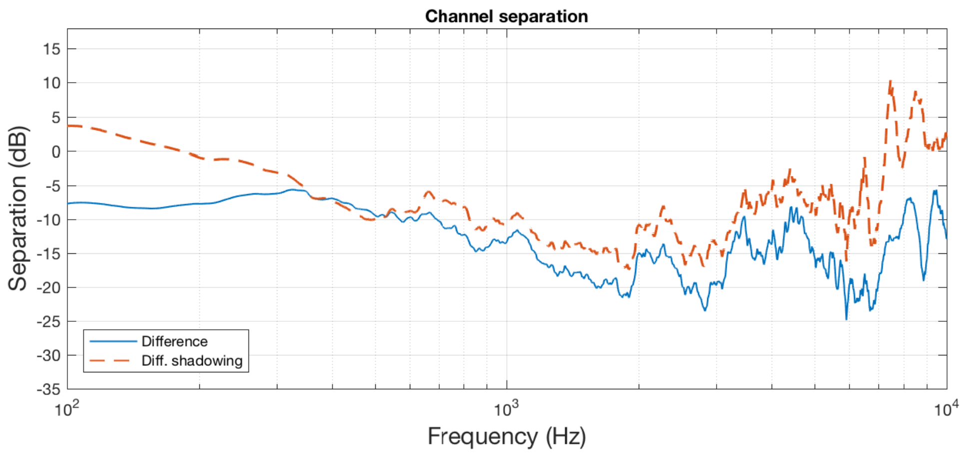


Figure 4.12: Channel separation for *front* system, *Test rig 3*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line, *Diff. shadowing*, shows the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

4.2 User study

The user study was performed in three steps: *User study pt.1, 2* and *3*. Two types of stimuli were chosen:

- short bursts of pink noise
- an anechoic recording of a female voice

The stimuli was placed in eight different virtual source positions evenly distributed along the circumference of a circle, as seen in figure 4.13:

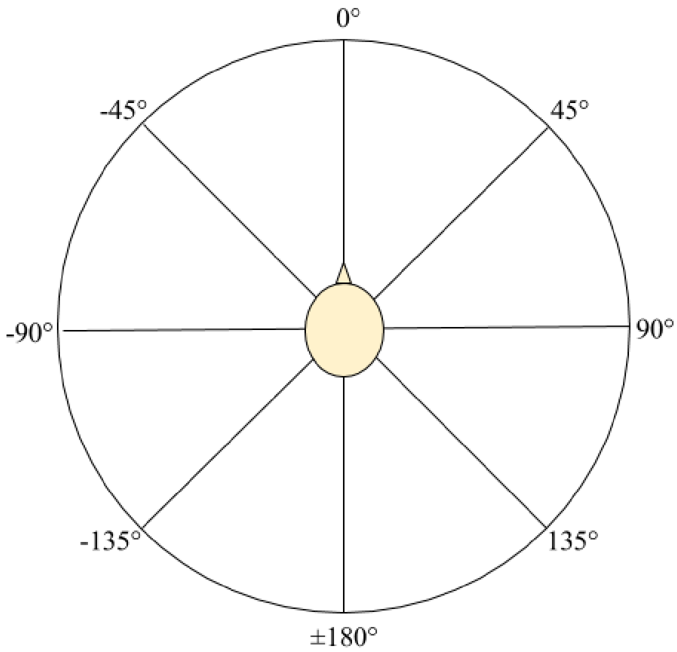


Figure 4.13: The virtual source positions used in the user studies.

All positions were tested twice for each stimuli and system resulting in a total of 64 played sounds per user study. The order was randomised for each participant and for each *User study pt.* The participants were introduced to each system before the test began with a short introduction to binaural playback systems. A binaural sound recording was played before each test started. The participants did not know which of the systems that was presenting the virtual source position. The user study design was tested in a pre-study with 3 participants.

The participants in the user studies had little or no experience with binaural playback systems and the age span was 29 to 54 years. A total of 14 individuals participated in the tests. The number of participants for each study was as follows:

- *User study pt.1*, 9 participants
- *User study pt.2*, 11 participants
- *User study pt.3*, 10 participants

4.2.1 Setup

All tests were conducted at Alpine Electronics using the custom built power amplifier. *User study pt.1* used *Test rig 1* consisting of four custom built loudspeakers and the setup was placed in the combined listening and conference room at Alpine's office. *User study pt.2* was conducted with *Test rig 2* inside a Volvo V40 at the garage of Alpine Electronics. The final study, *User study pt.3*, was conducted in the same room as *User study pt.1* with *Test rig 3*. In order to investigate the effects of using different set of custom built loudspeakers compared to the set used in *User study pt.1*.

4.2.2 Results

The results from the user study are presented in the following figures. The plots show the played position along the x-axis and the answers along the y-axis. A perfect system would have a diagonal ranging from (-180,180) up to (180,180). Answers located perpendicular to the diagonal is interpreted as front/back confusions. In figure 4.14 the results for the *back* system in setup *User study pt.1* are seen.

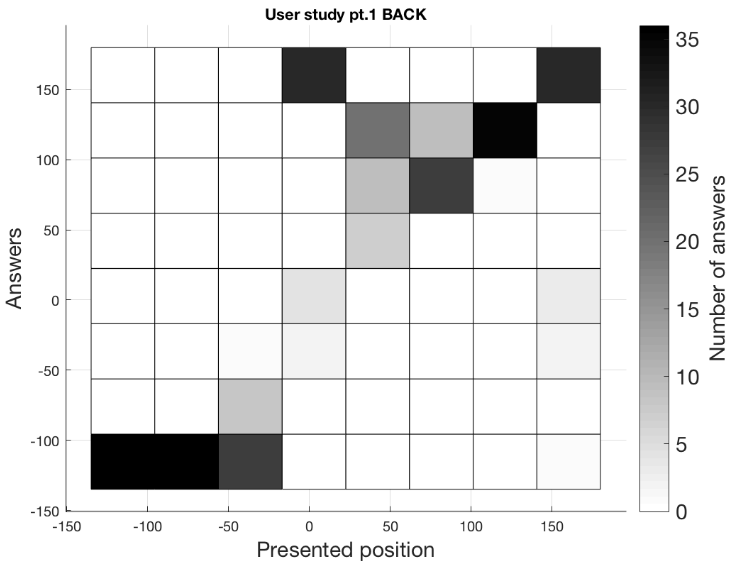


Figure 4.14: User study pt.1 results, back system.

As seen in the figure the source positions behind the listener has the highest hit rate of answers. The system seem to have slightly better reproduction at directions $+90°$ and $+134°$ compared to $-90°$ and $-134°$. For the front positions, $-45°$, $0°$ and $+45°$ the answers are mirrored towards the corresponding back positions i.e. front/back confusion was observed.

In figure 4.15 the results for the *back* system in setup *User study pt.2* are seen.

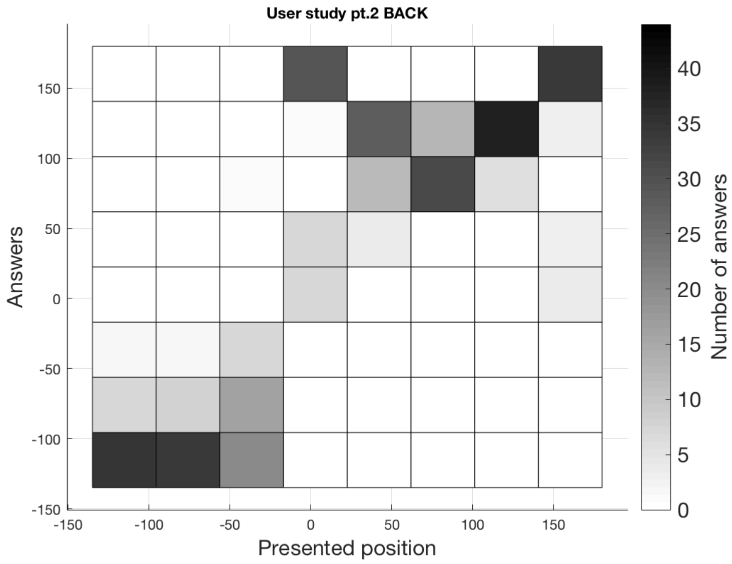


Figure 4.15: User study pt.2 results, back system.

This configuration displays a similar trend with more answers located in the back region. A slight improvement in localisation for the front positions as well as for $-90°$ compared to *User study pt.1* can be observed. As seen in the previous setup, there is still a tendency for better representation at $+90°$ and $+134°$ compared to $-90°$ and $-134°$.

In figure 4.16 the results for the *back* system in setup *User study pt.3* are seen.

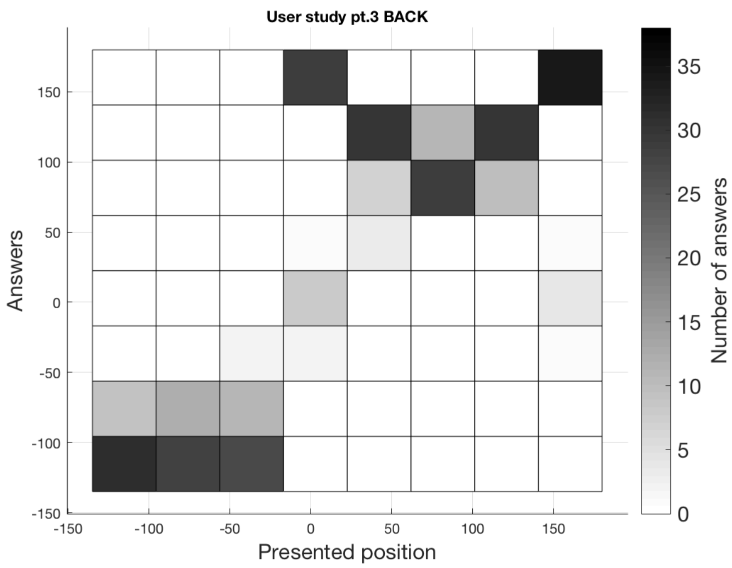


Figure 4.16: User study pt.3 results, back system.

Again, the same trend as for the previous setups is observed with the back positions acquiring a higher answer hit rate. Front/back confusions occur for the source positions presented in front of the listener where the virtual sources are perceived to be presented from behind. As for both the previous setups, these results also show a slightly better representation for positions at +90° and +134° compared to −90° and −134°.

The following figures show the results for the *front* system. In figure 4.17 the results for setup *User study pt.1* are seen.

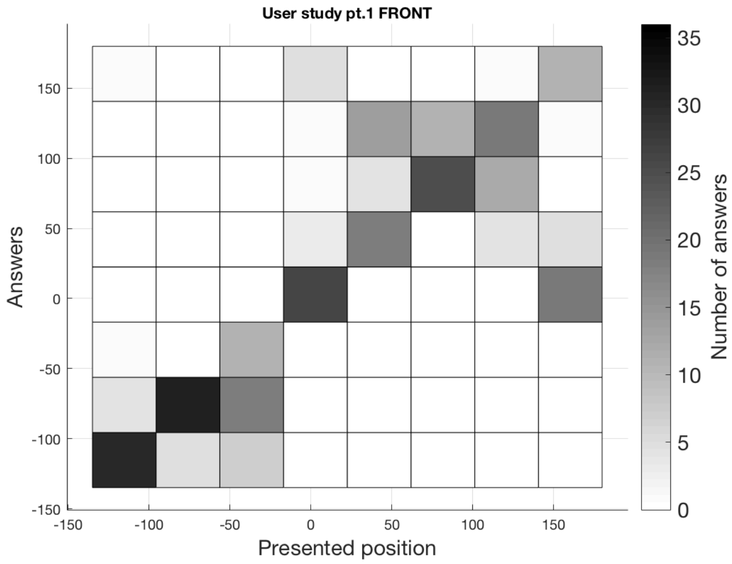


Figure 4.17: User study pt.1 results, front system.

In this setup, a clearer diagonal is shown, compared to the back system setups, with slightly blurred representations at positions +45°, +90°, +135° and +180° where more front/back confusions occur.

In figure 4.18 the results from *User study pt.2* are presented.

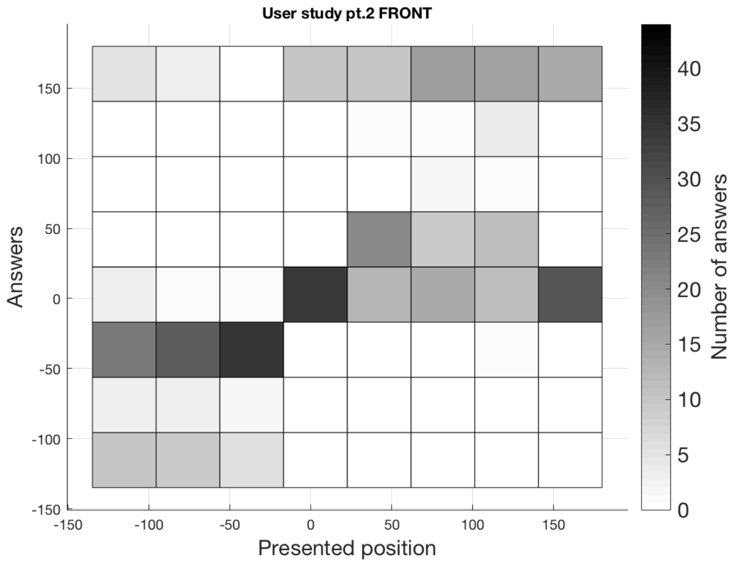


Figure 4.18: User study pt.2 results, front system.

In this test the answers are concentrated in the front stage of the tested positions. The presented angles +90°, +135°, −90° and −135° seemed to be difficult to place. Positive presented angles were perceived as placed behind or above the listener[2].

The results for *front* system in *User study pt.3* are presented in figure 4.19.

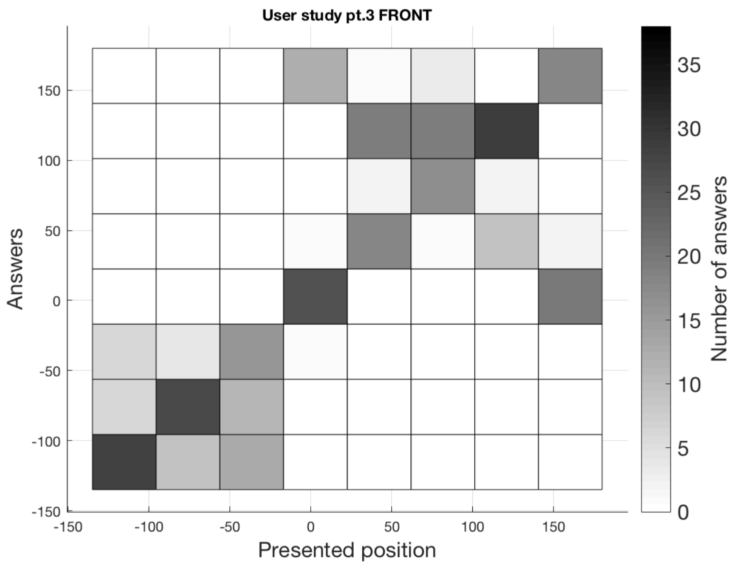


Figure 4.19: User study pt.3 results, front system.

As seen in the plot, the answers have a similar trend as seen in figure 4.17, where a hit rate of correct positions can be seen across the diagonal of the plot. However,

[2]Based on comments from the participants during the test.

the occurrence of front/back confusions are quite prominent for certain positions.

Since front/back confusion is a well documented phenomenon, corrections for front/back confusions were made in order to see if any other trends of each system can be seen. Each answer which resulted in a front/back confusion was mirrored to a corresponding position. These plots is found are *Appendix A*.

To further evaluate and compare each system against one another, the ratio q_i between the number of correct answers and total number of answers for each system was calculated, index i for back or front system. The ratios are presented in table 4.1, where $q_i = 1$ is a flawless system.

Table 4.1: Ratios of correct answers, q_i, for the different setups.

User study	q_b	q_f
pt.1	0.486	0.594
pt.2	0.466	0.349
pt.3	0.466	0.559

With this evaluation method, *Test rig 1* and *3* using the *front* system have the highest ratios. For *User study pt.2* the *front* system has poor results, as previously seen in figure 4.18. The *back* system performs relatively well in all three setups, with *Test rig 1* as the slightly better performing setup. The ratios for the corrected answers are seen in table 4.2.

Table 4.2: Ratios of correct answers, q_i, for the different setups using corrected data.

User study	$q_{b,c}$	$q_{f,c}$
pt.1	0.760	0.760
pt.2	0.702	0.577
pt.3	0.747	0.794

Apart from the overall expected improvement of the ratios, similar trends as for the uncorrected data are seen. However, the *front* system with *Test rig 3* has a slightly higher score than *Test rig 1*. These corrected ratios can be seen as a light indication of how the systems would perform if equalisation tuned to each participant was implemented.

4.2.3 Discussion

For the *back* system good channel separation was achieved in all three setups. Other observations from informal listening sessions throughout the filter tuning process as well as comments from the user studies point to the fact that the system was perceived as more robust. An explanation for this could be the head shadowing for the setup. Even with more detailed filters i.e. a low value for the number of bins in the moving average filter, the *back* setup had difficulties of placing virtual sources in front of the listener.

The *front* system was able to place virtual source positions more accurately, for both front and back positions, in setups placed in the listening room. For the setup inside the car, the system had problems of placing the sources, making the outcome of the user study a bit blurred. Many of the participants felt like the presented stimuli from the system was appearing above the head or coming from all different directions at the same time. A possible explanation for the less impressive performance could be the angling of the mounted loudspeaker drivers. The drivers mounted in the ceiling of the car were pointing more towards the floor in front of the listener than towards the listeners' ear positions compared to the setups in the listening room. Therefore the calculated filters are compensating for the off-axis frequency response, resulting in a filter which needs to be pushed harder in order to achieve a flat response. Further investigations regarding the possible effects of angling the *front* system would be of interest.

When the participants were asked about their experience from *Test rig 2* the virtual source positions were placed like a halo around the listener, with front positions appearing at the forehead of the listener. A possible reason for the limited spatial reproduction could be the rough estimation method in combination with a reverberant room. The created filters compensate for a semi anechoic environment, but the system is placed in a room with more reflective surfaces thus providing the listener with spatial cues which could be misleading when locating the virtual source positions. However, the system performs better in the listening room than inside the car, which suggests that the filter design method can be suitable for acoustically controlled rooms.

The results of the user study might have been biased by the fact the the loudspeaker positions were visible for the participants, thus influencing the perception towards the front or back positions e.g. by listening for which system that is actually playing instead of where the virtual source position is placed. Some of the participants commented on the fact that they "tuned in" on each system and thereby answered accordingly. This might also be an explanation for the "halo effect" previously mentioned. A different setup of "dummy speakers" or hidden speakers would have been preferable in retrospect.

Another factor that might have influenced the results is the fact that the systems were not equalised individually for each participant. This would improve the occurring front/back confusions.

The custom built loudspeakers for the *back* system used in *Test rig 2* and *3* had different drivers than the ones used in *Test rig 1*. The different directivity for the used loudspeakers might influence the resulting head shadowing for these setups.

5
Conclusion

With the use of a measurement system and a simple set of loudspeakers a binaural reproduction system can be implemented by using the presented filter design method and MATLAB code. Filters can be implemented using "stand alone" hardware such as digital signal processors, thereby simplifying installations. The field of 3D sound is presently growing along with the virtual reality market and the use of crosstalk cancellation technology can be implemented in various concepts. However, a constraining factor with the use of loudspeakers as a binaural reproduction system is the robustness. In car compartments the use of 3D sound can enable arbitrary placements of infotainment and warning sounds or create a totally different listening environment. The presented method can reproduce binaural audio inside a car and can be promising if the spatial resolution is not of great importance. As mentioned in the implementation chapter, the work towards the presented method included a lot of testing. Informal listening tests were conducted regularly by listening to binaural audio in all setups. These tests suggest that a virtual soundscape can be presented even in more acoustically challenging rooms when a user experience and listening effect is sought. The same binaural audio material was also presented to all participants during the user study and the comments were uniformly good for all setups in terms of user experience.

The space around the listener should be acoustically treated when pursuing an improved implementation of binaural reproduction systems inside cars. As shown in the results from the user study the room strongly effects the performance. Different headrest designs as well adding absorption to the main reflective surfaces would be interesting to investigate further.

Since the filters are compensating for the loudspeaker response, the choice of loudspeakers can affect the overall performance with the use of coarse filters. As previously presented, *Test rig 1* scored the highest ratio of correct answers for both the *back* and *front* systems. However, with corrected answers the vague trend ceases to appear so further testing would be needed to support this claim.

The rough filter design can be suitable for implementing a binaural reproduction system in a room with few reflecting surfaces close to the listening position. Using the *back* system as it is can be suitable to present binaural audio in source positions located behind and $\pm 90°$ of the listener.

Since the *back* system had difficulties of presenting sources in front of the listener, the filters would need to be improved e.g. by individual tuning. More precise filters could help tackle the challenges in a reverberant room, but with increased detail, the robustness will generally suffer. A way to improve the robustness is to use averages of different listening positions, as presented in reference [3]. Another way to increase robustness might be to use the *Least Mean Squares* method i.e. adaptive filters with respect to head movement. This can be implemented with the additional use of head tracking equipment and microphones. In the presence of such equipment, one might also use a number of static filters where each filter is representing a certain head position or rotation. Due to the poor performance of the *front* system implemented inside the car, another loudspeaker setup might be beneficial e.g. using an array of loudspeakers. A completely different filter design might be needed to improve the performance.

Speculations about what might cause the front/back confusions include the fact that the coarse filters might not compensate enough for the occurring reflections of the room. Furthermore, the colouring of each system, in addition to the participants seeing the actual loudspeaker placements caused the listeners to tune in to the system which would be more likely to emit the sounds. The comments from the user study suggest that the setups had different timbre and that some of the participants eventually figured out which system that was currently playing.

References

[1] A. Lundkvist, A. Nykänen, and R. Johnsson, "3d-sound in car compartments based on loudspeaker reproduction using crosstalk cancellation", in *130th Audio Engineering Society convention 2011*. Curran, 2011, vol. 1, pp. 203–213.

[2] F. Bellotti, R. Berta, A. D. Gloria, and M. Margarone, "Using 3d sound to improve the effectiveness of the advanced driver assistance systems", English, *Personal and Ubiquitous Computing*, vol. 6, no. 3, pp. 155–163, May 2002.

[3] J. Wang, Q. h. Ye, C. s. Zheng, and X. d. Li, "A robust algorithm for binaural audio reproduction using loudspeakers", in *2010 International Conference on Measuring Technology and Mechatronics Automation*, vol. 1, Mar. 2010, pp. 318–321.

[4] D. B. Ward and G. W. Elko, "Effect of loudspeaker position on the robustness of acoustic crosstalk cancellation", *IEEE Signal Processing Letters*, vol. 6, no. 5, pp. 106–108, May 1999, ISSN: 1070-9908.

[5] P. A. Nelson and S. J. Elliott, *Active control of sound*, English. London: Academic Press, 1992.

[6] T. McKelvey, "Project 2: Noise cancellation rev 1", Signals and systems, Chalmers university of technology, Tech. Rep., 2015.

[7] J. O. S. III. (2007). Introduction to digital filters, [Online]. Available: `https://ccrma.stanford.edu/~jos/filters/` (visited on 12/30/2016).

[8] P. Andersson, "Anc applications. block diagrams, x-lms implementation and the invertibility of a room ir.", Civil and Environmental Engineering, Applied Acoustics, Chalmers university of technology, Tech. Rep., 2003.

[9] S. W. Smith, *The scientist and engineer's guide to digital signal processing*, English. San Diego, Calif: California Technical Publ, 1997.

[10] Mathworks. (2016). Curve fitting toolbox, user's guide, [Online]. Available: `https://se.mathworks.com/help/curvefit/index.html` (visited on 12/30/2016).

A

Appendix A

A.1 In-vehicle measurements

This section is intended as a clarification for the measurement method and setups. The measurements were performed inside a Volvo V40 situated at the garage of Alpine Electronics. External noise sources were passing vehicles, ventilation noise and thumps from a nearby gym. The recordings were averaged with 6 averages and timed when the background noise was as low as possible. Due to practical reasons, the back right passenger door was fully open during the measurements. In table A.1 the measured positions are declared. Reference points were marked on the neck of the *HATS* with 1 cm steps for rotating the head. All measurements were using the same script and program settings.

Table A.1: Measurement placement 1 setup for the *HATS*.

Name	**Position**
01	centred
02	1 cm turn left
03	2 cm turn left
04	3 cm turn left
05	4 cm turn left
06	1 cm turn right
07	2 cm turn right
08	3 cm turn right
09	4 cm turn right

To verify that the measurement scripts were accurate additional measurements were performed using *REW*. The following positions were tested: 01, 03, 05, 07 and 09. The data acquired using *REW* were then compared with the data captured using the measurement scripts. Both methods gave the same results.

A.1.1 Placement setup 1

The loudspeakers positions were based on the outcome of informal testing of crosstalk performance and head shadowing in *Test rig 0*. The mounting was measured by sitting in the driving seat and marking reference points in the ceiling and side wall of the car door. These reference points were used to estimate the positioning of the dummy head. Speakers were mounted one at a time to verify the height and resulting driver position. The plate with mounted speakers was then placed at the headrest and secured using straps. Due to no fixation of the *HATS*, it might have moved slightly between each new rotation. However, the distance to the ears from the speakers were controlled when the dummy was in the ideal listening position. The distance to each microphone seemed to be correct, in terms of the expected accuracy of a simple measuring tape.

A.1.2 Placement setup 2

To further investigate the robustness of the filters, additional positions of the *HATS* were measured. Instead of rotating the head the whole body was slanted to the left and to the right. Also, the *HATS* was slanted forward towards the steering wheel. To determine the dummy head position, the angle of the slanting was measured and estimated. The same procedure as with placement setup 1 was used; a piece of tape marked with a scale was put in the ceiling at the approximate centre of the dummy head for reference. Due to the difficulties of placing the *HATS* and measuring its position, the following tests are to be seen as informal and as mentioned earlier, as a guideline to see how robust the filters are in theory. The target positions are described in table A.2. However, the actual positions can be considered to be a bit arbitrary. The distance to the ears from the edge of the drivers are stated as a reference point. They seem to be fairly symmetrical from the approximated measurements made. For the slanting towards the steering wheel, position 30, a cylindrical cardboard piece was placed between the seating and the back of the dummy.

Table A.2: Measuring placement setup 2 for the *HATS*.

Name	**Position**
10	6 cm left
20	6 cm right
30	13,5 cm front

The results were analysed during the process of the project and the following was found:

- The *back* system was generally more robust in all movement directions. However, the systems' spatial reproduction was less accurate probably due to the decreasing crosstalk performance with increasing distance.
- The *front* system was very fragile against head movement. Slightest movement gave huge differences in crosstalk performance.

A.2 System performance and channel separation

In the following section the plots for the test rigs used in the user study are shown. Each rig has 4 plots starting with two plots for the *back* system and followed by two plots for the *front* system.

A.2.1 Test rig 1

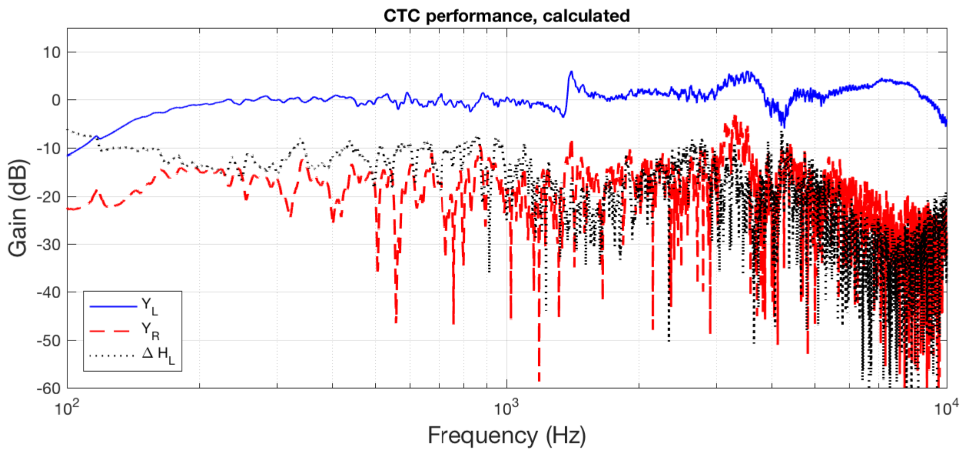


Figure A.1: Frequency response with crosstalk performance for *back* system, *Test rig 1*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

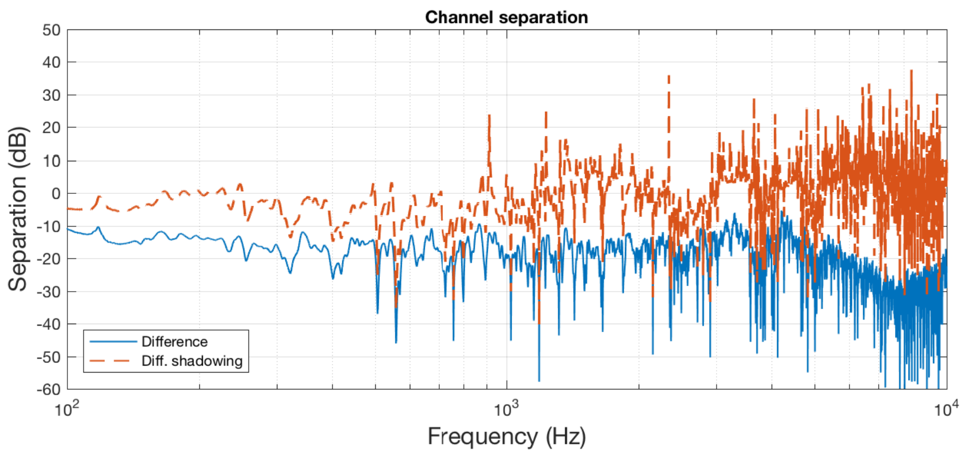


Figure A.2: Channel separation for *back* system, *Test rig 1*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line *Diff. shadowing* is the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

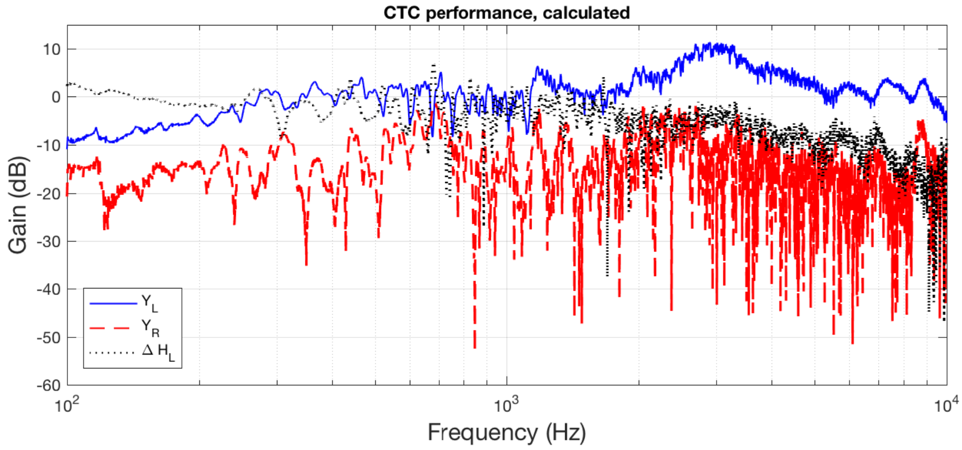


Figure A.3: Frequency response with crosstalk performance for *front* system, *Test rig 1*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

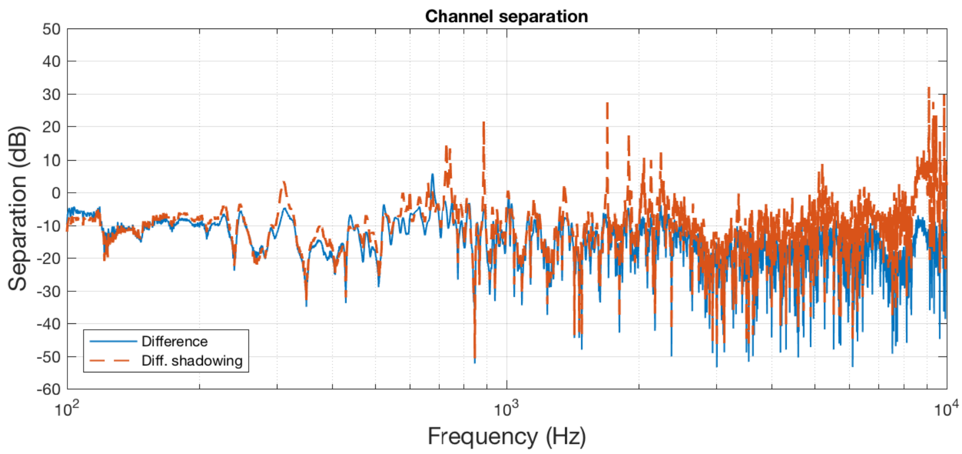


Figure A.4: Channel separation for *front* system, *Test rig 1*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line *Diff. shadowing* is the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

A.2.2 Test rig 2

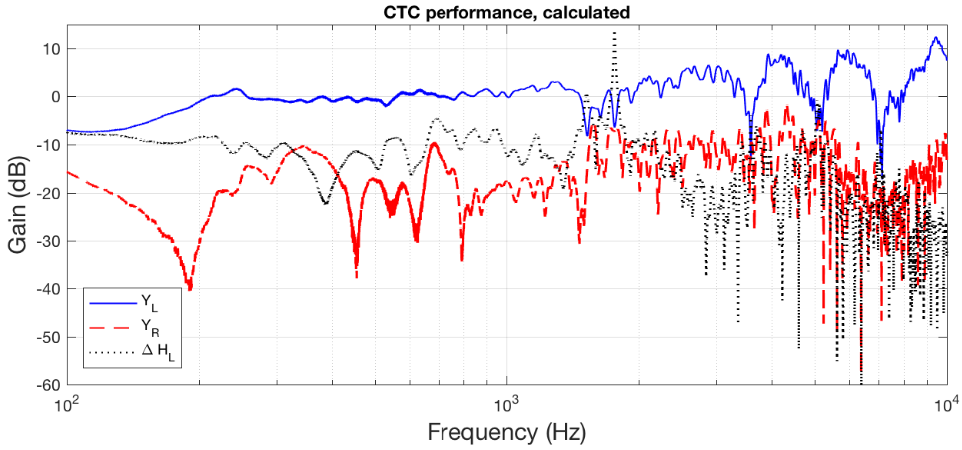


Figure A.5: Frequency response with crosstalk performance for *back* system, *Test rig 2*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

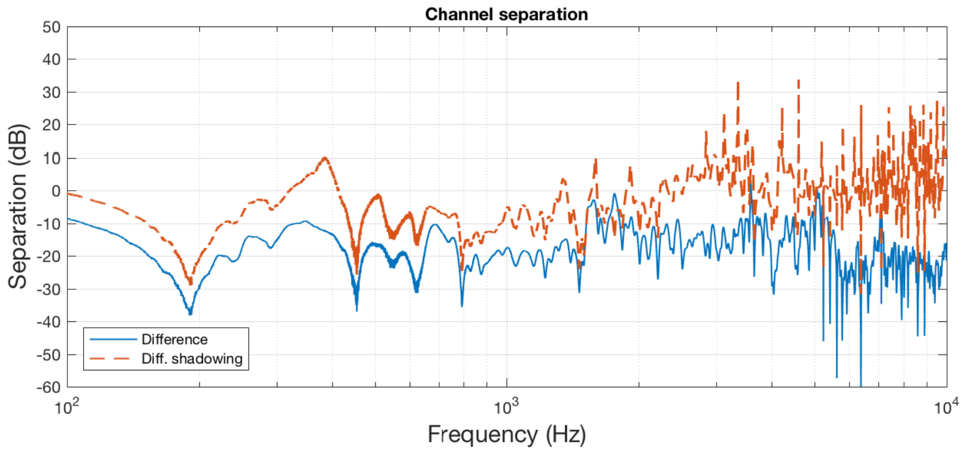


Figure A.6: Channel separation for *back* system, *Test rig 2*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line *Diff. shadowing* is the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

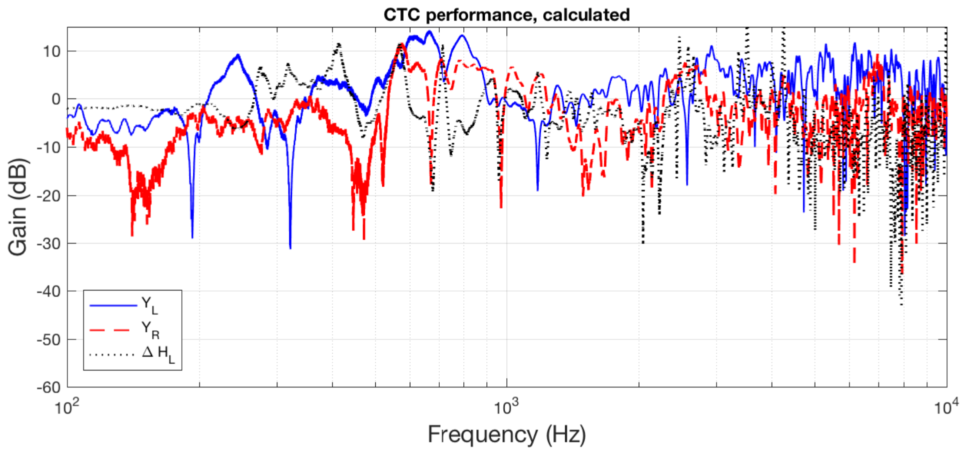


Figure A.7: Frequency response with crosstalk performance for *front* system, *Test rig 2*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

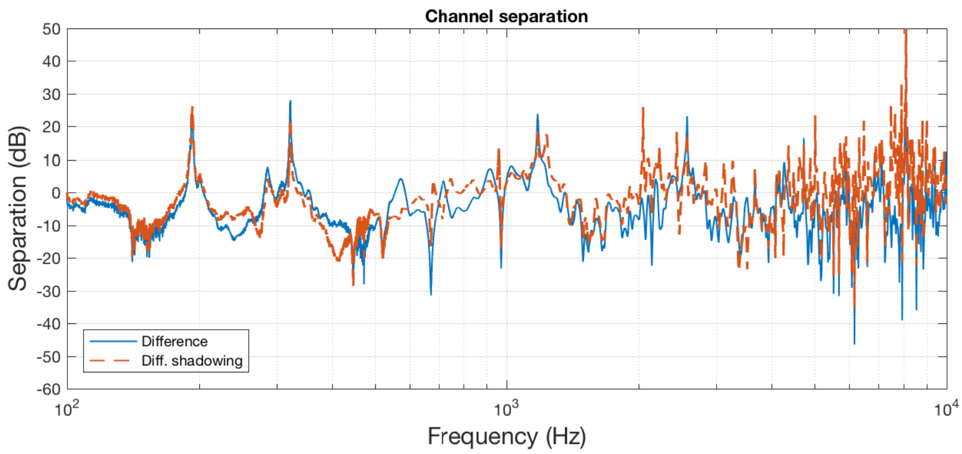


Figure A.8: Channel separation for *front* system, *Test rig 2*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line *Diff. shadowing* is the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

A.2.3 Test rig 3

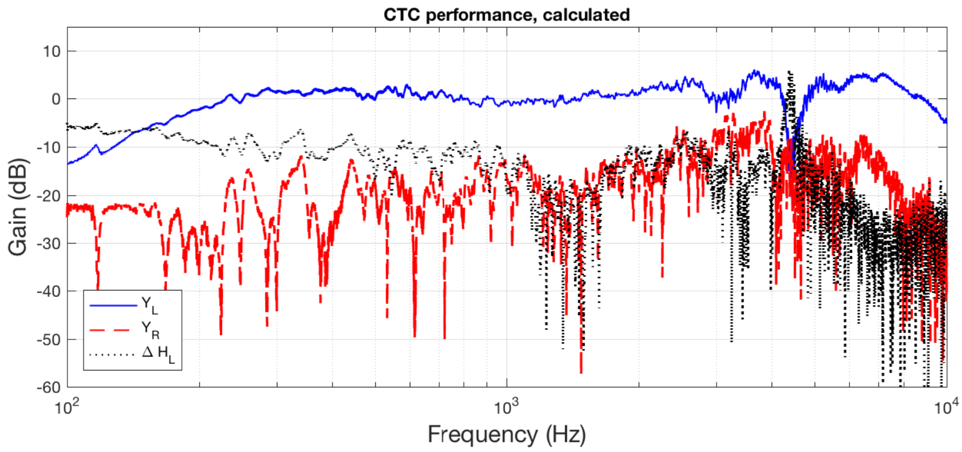


Figure A.9: Frequency response with crosstalk performance for *back* system, *Test rig 3*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

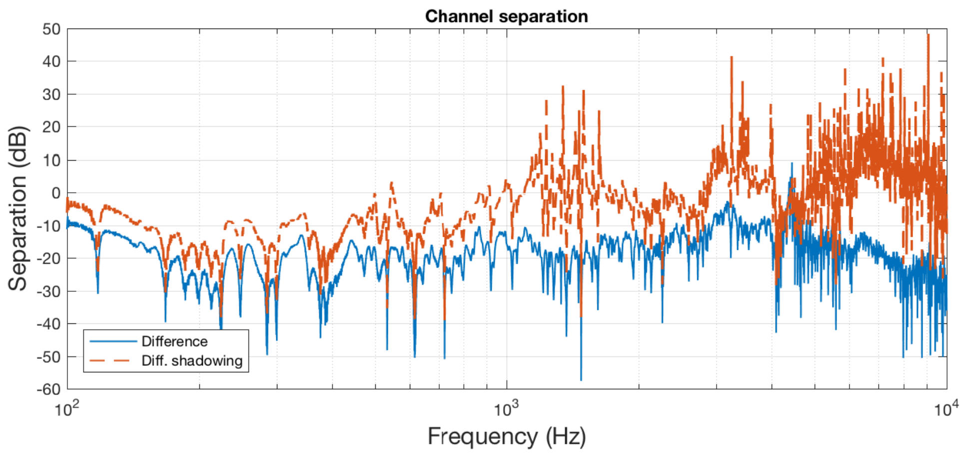


Figure A.10: Channel separation for *back* system, *Test rig 3*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line *Diff. shadowing* is the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

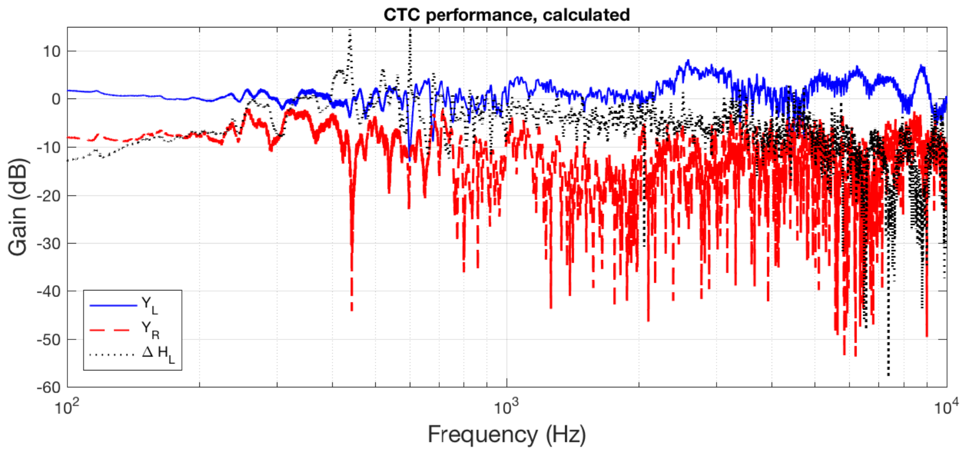


Figure A.11: Frequency response with crosstalk performance for *front* system, *Test rig 3*. Y_L and Y_R denote the illuminated and shadowed ear respectively. ΔH_L is the head shadowing.

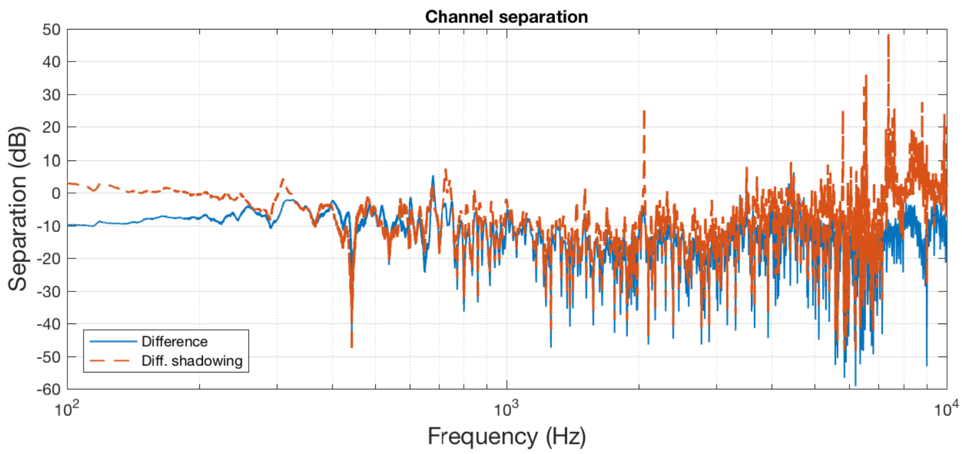


Figure A.12: Channel separation for *front* system, *Test rig 3*. The solid line shows the *difference* i.e. channel separation between receiving ears. The dashed line *Diff. shadowing* is the channel separation with head shadowing effects subtracted i.e. the filters contribution to channel separation.

A.3 Corrected user study answers

The following section declares the corrected user study results.

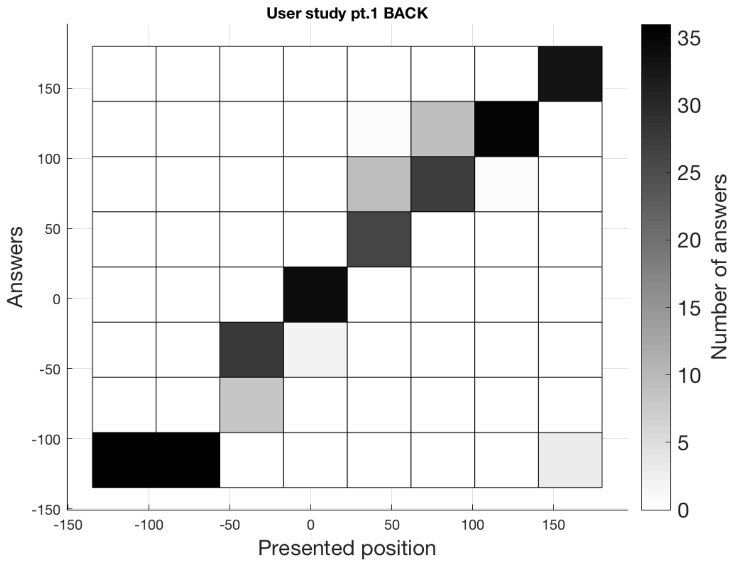


Figure A.13: User study pt.1, corrected front/back answers, back system.

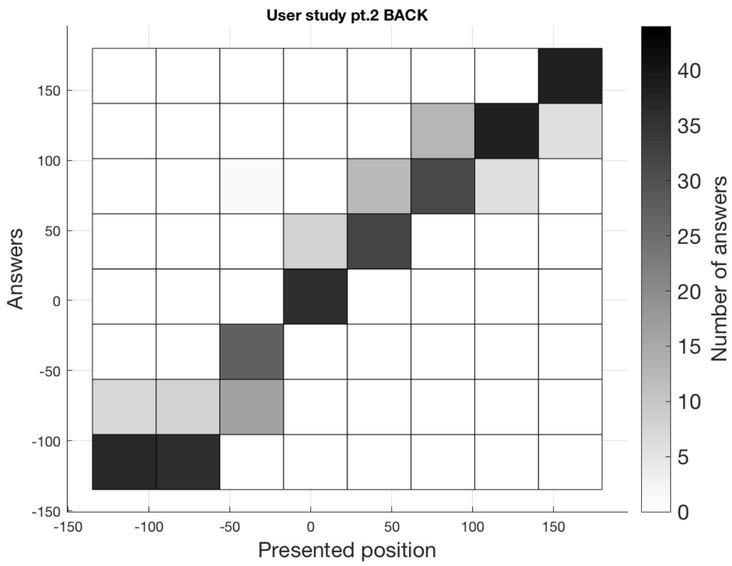


Figure A.14: User study pt.2, corrected front/back answers, back system.

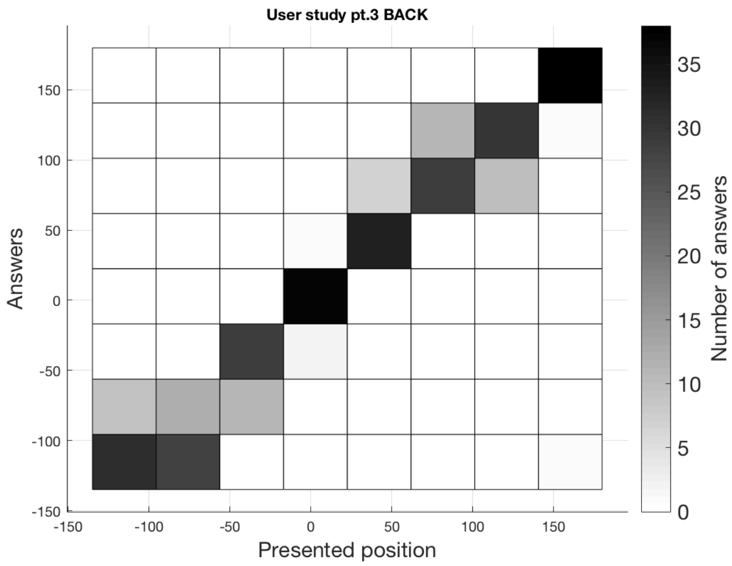


Figure A.15: User study pt.3, corrected front/back answers, back system.

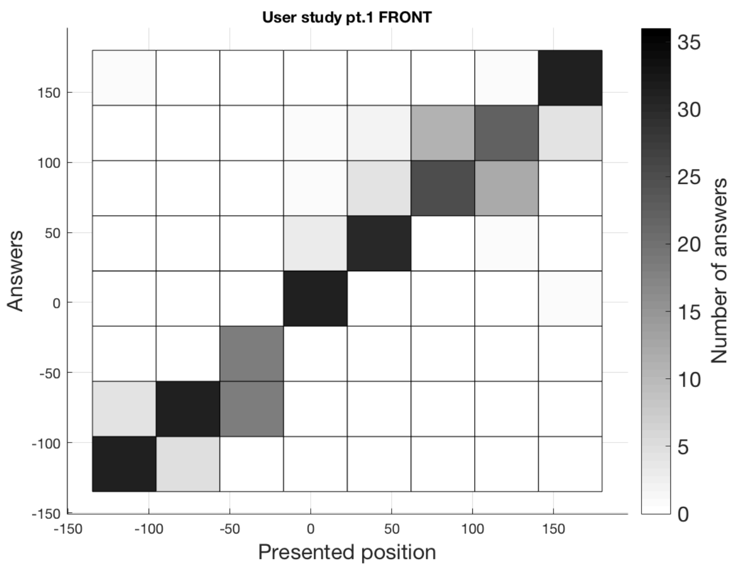


Figure A.16: User study pt.1, corrected front/back answers, front system.

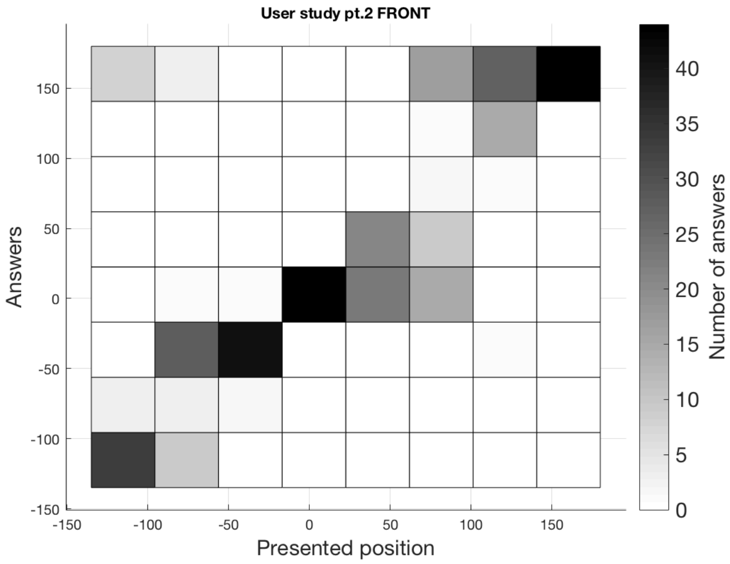


Figure A.17: User study pt.2, corrected front/back answers, front system.

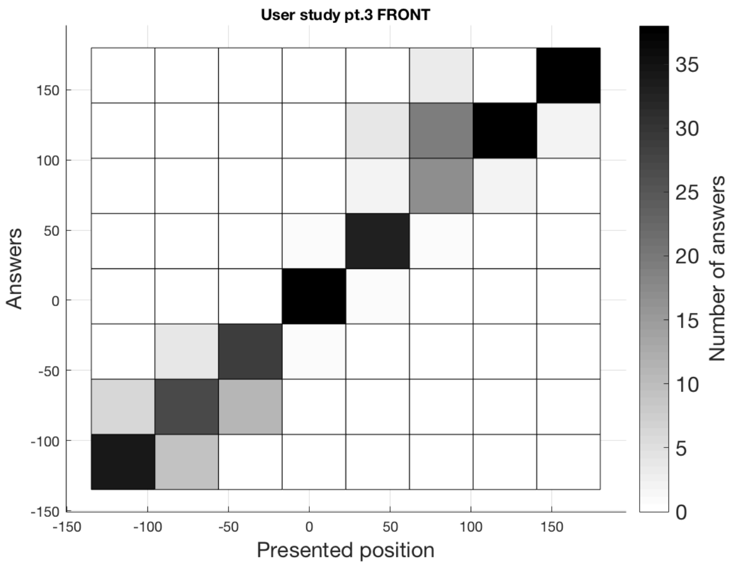


Figure A.18: User study pt.3, corrected front/back answers, front system.

A.4 Frequency response of used loudspeakers

In this section the frequency response of the used loudspeakers are presented. All measurements were performed using *REW* except for *front* system in *User study pt.2* and *3* where the software *Audio Precision* was used.

A.4.1 Loudspeakers used in user study pt.1

In figure A.19 the frequency response plots for the head rest- and the ceiling loudspeakers are presented.

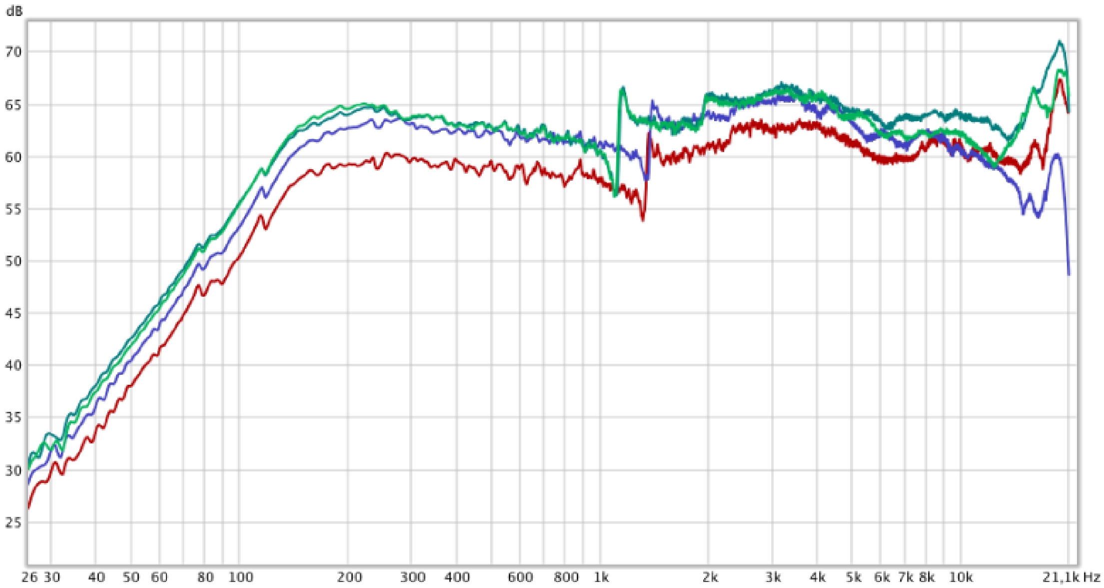


Figure A.19: Frequency response of custom built loudspeakers used in *User study pt.1*. Red line denote right loudspeaker for *back* system; blue line denote the left loudspeaker for *back* system; dark green line denote the left loudspeaker for *front* system and light green line denote the right loudspeaker for *front* system.

A.4.2 Loudspeakers used in user study pt.2 and pt.3

In figures A.20 and A.21 the frequency response plots for the head rest- and the ceiling loudspeakers are presented. These loudspeakers were used in both *User study pt.2* and *3*.

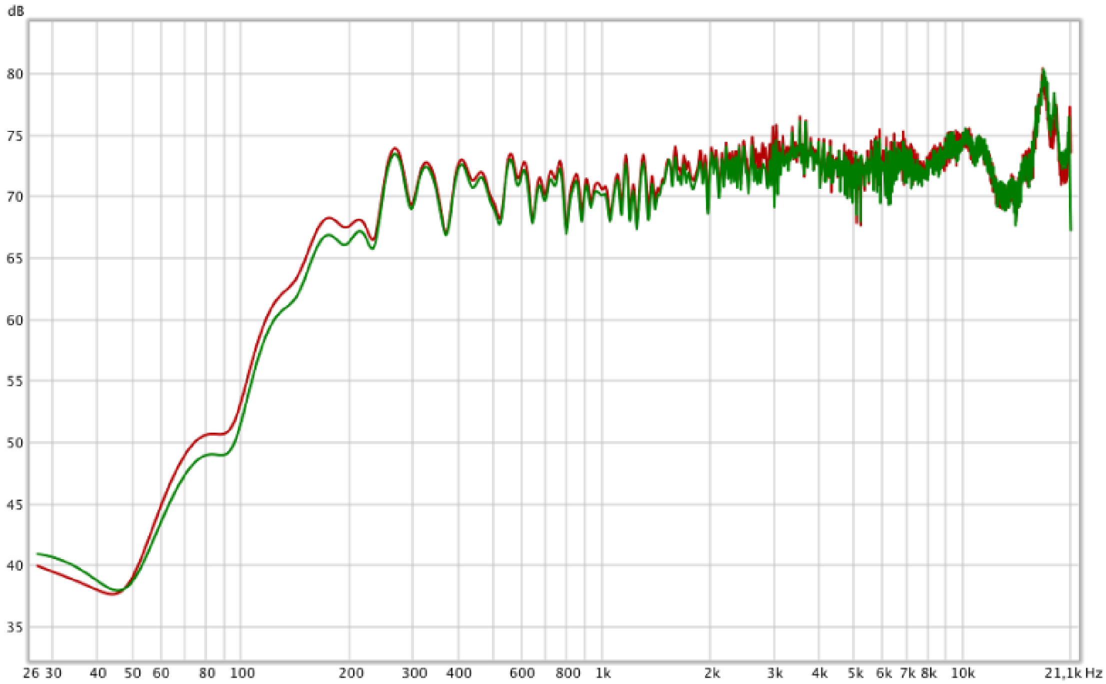


Figure A.20: Frequency response of custom built loudspeakers used in *User study pt.2* and *3*, *back* system. Red line denote the right loudspeaker and green line denote the left loudspeaker.

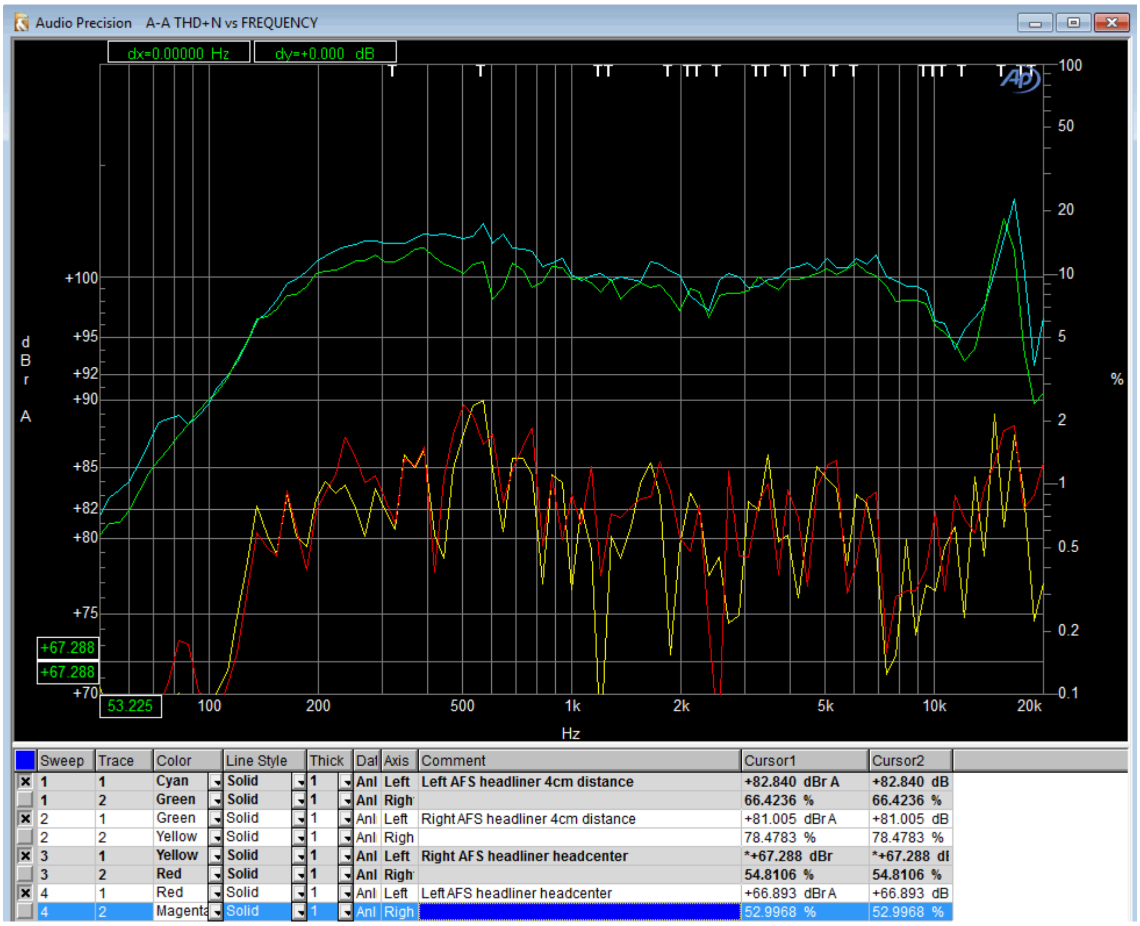


Figure A.21: Frequency response of custom built loudspeakers used in *User study pt.2* and *3*, *front* system. Cyan line denote the left loudspeaker and green line denote the right loudspeaker. The red and yellow line denotes distortion for the right and left loudspeaker respectively.

B
Appendix B

In this appendix the MATLAB code for the created scripts that were used are presented.

B.1 Measurement scripts

The following functions were used for measuring the propagation paths using a sine sweep as excitation signal. The code was used for recordings with a sampling frequency, f_s, of 48 kHz.

B.1.1 ir_capture.m

```
%% "ir_capture", measure IR, MSc-thesis Jonas Karlberg 2016
function [H,h,y]=ir_capture(x,nBits,fs,nChannels,IDin,IDout,stop,avg)
%% play and record sound

stoptime=(length(x)/fs);              %sets the length of recording

recObj = audiorecorder(fs,nBits,nChannels,IDin);        %sample rate, bitresolution, mono/stereo, input channel

y=zeros(length(x),avg*2);
index=1;

disp('Measuring.')
for loop=1:avg;
    playObj = audioplayer(x,fs,nBits,IDout);

    play(playObj)                                       %play stimuli

    recordblocking(recObj, stoptime);                   %start recording

    y_temp = getaudiodata(recObj);                          %store captured data in y
    y(:,index)=y_temp(:,1);
    y(:,index+1)=y_temp(:,2);

    index=index+2;
    clear playObj;
end
disp('End of Recording.');
```

```


%% estimate TFs i.e. deconvolve recorded signal

H=zeros(length(y(:,1)),avg);
h=zeros(length(y(:,1)),avg);
h_avg=zeros(length(y(:,1)),1);
Wn=2*60/fs;                    % cut off, 60 Hz (set this according to the sweep signal)
[b,a]=butter(6,Wn,'high');     %create hp-filter

index=1;
for loop=1:avg
    H(:,loop) = fft(y(:,index),length(y))./fft(y(:,(index+1)),length(y));    % channel 1: microphone, channel 2: fed back

    h(:,loop) = real(ifft(H(:,loop)));              %IR of measured system
    index=index+2;

    h_avg=h_avg+h(:,loop);
end

h=h_avg./avg;

h=filter(b,a,h);

stop=stop*fs;                                   %sets cut out length of IR in samples (end sample)
peak=find(h==max(h));                                   %finds IR peak

%%
minValue = 1;                                           %start-point for IR extraction
maxValue = stop;                                    %end-point for IR extraction

h = h(minValue:maxValue);                               %cuts IR

n=length(h);                                            %length of cut IR

if peak >=90
    fadein= [zeros(95,1); hann(round(20)); zeros(95,1)];    %create Hanning filter fade-in with 210 samples
else if peak < 90 && peak > 20
    fadein= [zeros(17,1); hann(round(18)); zeros(17,1)];    %create Hanning filter fade-in with 52 samples

else if peak < 20 && peak > 14

        fadein= [zeros(12,1); hann(round(12)); zeros(12,1)];    %create Hanning filter fade-in with 36 samples

else
      fadein= ones(28,1);                  %create "all pass" fade-in with 28 samples
```

```
        end
        end
end

fadeout= hann(2000);
fade=[fadein( 1:end/2 ); fadeout(end/2+1:end)];
window = [ fadein( 1:end/2 ); ones( n - length(fade), 1 ); fadeout(end
    /2+1:end) ];   %creates window

h = h.*window;                                      %applies window
    to cut IR

%% compute FFTs

H=fft(h);

return
```

B.1.2 create_stimuli.m

```
%% stimuli creation, MSc-thesis Jonas Karlberg 2016

function[sweep]=create_stimuli(fs,stop)

f0=20;                    %start frequency of sweep
f1=20000;                %end frequency of sweep
t=0:1/fs:stop;           %time vector

y=chirp(t,f0,stop,f1,'logarithmic');      %generate sweep

sweep=[.5*y.'; zeros(10000,1)];              %scale magnitude and add
    silence


n=length(sweep);
fadein= [zeros(40,1); hann(round(16000)); zeros(40,1)];
fadeout= hann(8000);
fade=[fadein( 1:end/2 ); fadeout(end/2+1:end)];
window = [ fadein( 1:end/2 ); ones( n - length(fade), 1 ); fadeout(end
    /2+1:end) ];   %creates window
sweep=sweep.*window;
sweep= [sweep, sweep];                     %for stereo signal

return
```

B.1.3 cutIR2.m

```
%% cut IR to desired length, MSc-thesis Jonas Karlberg 2016
function [h]=cutIR2(h,cut)

h1=h(1:cut,1);                      %cut out IR

n=length(h1);

```

```
fadeout= hann(n/2);
window = [ ones( n - length(fadeout)/2-30, 1 ); fadeout(end/2+1:end); zeros(30,1) ];   %creates window (only fadeout)

h1 = h1.*window;                                              %applies window to cut IR

h=h1;
```

B.1.4 windctIR.m

```
%% fix IR of cross-talk, Msc-thesis Jonas Karlberg 2016

function [h]=windctIR(h)

n=length(h);
fadein= [zeros(40,1); hann(round(40)); zeros(40,1)];
fadeout= hann(2000);
fade=[fadein( 1:end/2 ); fadeout(end/2+1:end)];

window = [ fadein( 1:end/2 ); ones( n - length(fade), 1 ); fadeout(end/2+1:end) ];   %creates window

h = h.*window;                                              %applies window to cut IR
```

B.2 Calculation of CTC-filters

B.2.1 ctc_calc_smooth5.m

```
%% Calculation of CTC filters, MSc-thesis Jonas Karlberg 2016
%%% calculates ideal CTC filters and regularises afterwards using:
%%% smooth, rceps
%%% A --> min phase --> C --> smooth --> limit magnitude --> smooth pahse --> combine magnitude and phase --> return filter (double sided)
function [C_LL,C_LR,C_RL,C_RR] = ctc_calc_smooth5(smoothing,ftrans,FRFs,fs,f_c,allpass)

N1=smoothing(1,1);
N2=smoothing(1,2);
N3=smoothing(1,3);
N4=smoothing(1,4);

H_LL=FRFs(:,1);
H_LR=FRFs(:,2);
H_RL=FRFs(:,3);
H_RR=FRFs(:,4);

%%% denominator for CTC-matrix
A=(H_RR.*H_LL-H_RL.*H_LR);

a=real(ifft(A));
[~, ay]=rceps(a);          %create min-phase response
```

```
A_mp=fft(ay);

%%% ideal CTC filters with min phase denominator
C_LL=H_RR./A_mp;
C_RL=-H_RL./A_mp;
C_LR=-H_LR./A_mp;
C_RR=H_LL./A_mp;

%%% apply high pass filter
c_ll=real(ifft(C_LL));
c_rl=real(ifft(C_RL));
c_lr=real(ifft(C_LR));
c_rr=real(ifft(C_RR));

Wn=2*100/fs;                    %cut off frequency 100 Hz
[b,a]=butter(6,Wn,'high');      %create butterwoth highpass filter

c_ll_hp=filter(b,a,c_ll);       %filter response
c_lr_hp=filter(b,a,c_lr);
c_rl_hp=filter(b,a,c_rl);
c_rr_hp=filter(b,a,c_rr);

%%% apply low pass filter
Wn=2*f_c/fs;                  %cut off frequency 12000 Hz
[b,a]=butter(24,Wn,'low');      %create butterwoth lowpass filter

c_ll_bp=filter(b,a,c_ll_hp);      %filter response
c_lr_bp=filter(b,a,c_lr_hp);
c_rl_bp=filter(b,a,c_rl_hp);
c_rr_bp=filter(b,a,c_rr_hp);

%%% apply window
n=length(ay);

fadein= hann(10);
fadeout= hann(n/2);

fade=[fadein( 1:end/2 ); fadeout(end/2+1:end)];

window = [ fadein( 1:end/2 ); ones( n - length(fade), 1 ); fadeout(end
    /2+1:end) ];

c_ll_wndw = c_ll_bp.*window;
c_lr_wndw = c_lr_bp.*window;
c_rl_wndw = c_rl_bp.*window;
c_rr_wndw = c_rr_bp.*window;

C_LL=fft(c_ll_wndw);
C_LR=fft(c_lr_wndw);
C_RL=fft(c_rl_wndw);
C_RR=fft(c_rr_wndw);

%%% split magnitude and phase for the CTC-filters
C_LL_abs=abs(C_LL);
C_LR_abs=abs(C_LR);
C_RL_abs=abs(C_RL);
```

```
C_RR_abs=abs(C_RR);

C_LL_angle=angle(C_LL);
C_LR_angle=angle(C_LR);
C_RL_angle=angle(C_RL);
C_RR_angle=angle(C_RR);

%%% smooth response of filters using N1 and N2 (split smoothing)

C_LL_sm1=smooth(C_LL_abs,N1);
C_LR_sm1=smooth(C_LR_abs,N1);
C_RL_sm1=smooth(C_RL_abs,N1);
C_RR_sm1=smooth(C_RR_abs,N1);

C_LL_sm2=smooth(C_LL_abs,N2);
C_LR_sm2=smooth(C_LR_abs,N2);
C_RL_sm2=smooth(C_RL_abs,N2);
C_RR_sm2=smooth(C_RR_abs,N2);

n1=length(C_LL);
f1=linspace(0,fs,n1);

ftrans_rev=fs-ftrans;      %for upper part of spectrum

indsfSmoothLF=find(f1<=ftrans).';
indsfSmoothHF=find(f1>=ftrans & f1<=ftrans_rev).';
indsfSmoothLFrev=find(f1>=ftrans_rev).';

C_LL_sm=[C_LL_sm1(indsfSmoothLF);C_LL_sm2(indsfSmoothHF);C_LL_sm1(indsfSmoothLFrev)];
C_LR_sm=[C_LR_sm1(indsfSmoothLF);C_LR_sm2(indsfSmoothHF);C_LR_sm1(indsfSmoothLFrev)];
C_RL_sm=[C_RL_sm1(indsfSmoothLF);C_RL_sm2(indsfSmoothHF);C_RL_sm1(indsfSmoothLFrev)];
C_RR_sm=[C_RR_sm1(indsfSmoothLF);C_RR_sm2(indsfSmoothHF);C_RR_sm1(indsfSmoothLFrev)];

%%% fix magnitude (ALL-PASS filter OR HIGH SHELF over f_2)


indsfshelfLF=find(f1<=f_c).';                        %indicies below filter limit
indsfshelfHF=find(f1>=f_c & f1<=(fs-f_c)).';     %indicies above filter limit
indsfshelfLFrev=find(f1>=(fs-f_c)).';             %mirrored indicies

if allpass==1   %% all-pass filter

          C_LL_allpass(indsfshelfHF,1)=1;
          C_LR_allpass(indsfshelfHF,1)=1;
          C_RL_allpass(indsfshelfHF,1)=1;
          C_RR_allpass(indsfshelfHF,1)=1;

          HF_LL=C_LL_allpass(indsfshelfHF);
          HF_LR=C_LR_allpass(indsfshelfHF);
          HF_RL=C_RL_allpass(indsfshelfHF);
```

```
          HF_RR=C_RR_allpass(indsfshelfHF);

      else if allpass==0 %high shelf filter

            index=indsfshelfLF(end);        %index of the last value

            C_LL_shelf(indsfshelfHF,1)=C_LL_sm(index);
            C_LR_shelf(indsfshelfHF,1)=C_LR_sm(index);
            C_RL_shelf(indsfshelfHF,1)=C_RL_sm(index);
            C_RR_shelf(indsfshelfHF,1)=C_RR_sm(index);

            HF_LL=C_LL_shelf(indsfshelfHF);
            HF_LR=C_LR_shelf(indsfshelfHF);
            HF_RL=C_RL_shelf(indsfshelfHF);
            HF_RR=C_RR_shelf(indsfshelfHF);
        else
            HF_LL=C_LL_sm(indsfshelfHF);
            HF_LR=C_LR_sm(indsfshelfHF);
            HF_RL=C_RL_sm(indsfshelfHF);
            HF_RR=C_RR_sm(indsfshelfHF);
        end
end

[C_LL_sm]=[C_LL_sm(indsfshelfLF);HF_LL;C_LL_sm(indsfshelfLFrev)];
[C_LR_sm]=[C_LR_sm(indsfshelfLF);HF_LR;C_LR_sm(indsfshelfLFrev)];
[C_RL_sm]=[C_RL_sm(indsfshelfLF);HF_RL;C_RL_sm(indsfshelfLFrev)];
[C_RR_sm]=[C_RR_sm(indsfshelfLF);HF_RR;C_RR_sm(indsfshelfLFrev)];

%%% smooth overlaps from split smoothing & allpass/shelf operation

[C_LL_sm]=smooth(C_LL_sm,N1);
[C_LR_sm]=smooth(C_LR_sm,N1);
[C_RL_sm]=smooth(C_RL_sm,N1);
[C_RR_sm]=smooth(C_RR_sm,N1);

%%% split-smooth phase

% Smooth phase with N3 point moving average
C_LL_angle1=smooth(unwrap(C_LL_angle),N3);
C_LR_angle1=smooth(unwrap(C_LR_angle),N3);
C_RL_angle1=smooth(unwrap(C_RL_angle),N3);
C_RR_angle1=smooth(unwrap(C_RR_angle),N3);

% Smooth phase with N4 pont moving average
C_LL_angle2=smooth(unwrap(C_LL_angle),N4);
C_LR_angle2=smooth(unwrap(C_LR_angle),N4);
C_RL_angle2=smooth(unwrap(C_RL_angle),N4);
C_RR_angle2=smooth(unwrap(C_RR_angle),N4);

            HF_LL_phase=C_LL_angle2(indsfshelfHF);
            HF_LR_phase=C_LR_angle2(indsfshelfHF);
            HF_RL_phase=C_RL_angle2(indsfshelfHF);
            HF_RR_phase=C_RR_angle2(indsfshelfHF);

% combine smoothened-phases
[C_LL_angle_sm]=[C_LL_angle1(indsfshelfLF);HF_LL_phase;C_LL_angle1(
```

```
    indsfshelfLFrev)];
[C_LR_angle_sm]=[C_LR_angle1(indsfshelfLF);HF_LR_phase;C_LR_angle1(
    indsfshelfLFrev)];
[C_RL_angle_sm]=[C_RL_angle1(indsfshelfLF);HF_RL_phase;C_RL_angle1(
    indsfshelfLFrev)];
[C_RR_angle_sm]=[C_RR_angle1(indsfshelfLF);HF_RR_phase;C_RR_angle1(
    indsfshelfLFrev)];

C_LL_angle_sm=smooth((C_LL_angle_sm),200);
C_LR_angle_sm=smooth((C_LR_angle_sm),200);
C_RL_angle_sm=smooth((C_RL_angle_sm),200);
C_RR_angle_sm=smooth((C_RR_angle_sm),200);

%%% reconstruct phase

C_LL=C_LL_sm.*exp(1i.*C_LL_angle_sm);
C_LR=C_LR_sm.*exp(1i.*C_LR_angle_sm);
C_RL=C_RL_sm.*exp(1i.*C_RL_angle_sm);
C_RR=C_RR_sm.*exp(1i.*C_RR_angle_sm);

return
```

B.2.2 limitFRF.m

```
%% limit and fix frequency response, MSc-thesis Jonas Karlberg 2016

function [Cout]=limitFRF(Cin,fs,f_1,f_2)

n=length(Cin);

%%% create hp and lp filter

freq = [0 2*f_1/fs 2*f_2/fs 1];       %set stop and pass-bands
a = [0 0 1 1];

hp=firls(254,freq,a);          %design linear phase hp-filter

center=ceil(length(hp)/2);  %finds center of the created filter
c=zeros(1,length(hp));

c(center)=1;

lp=c-hp;                       %creates lp-filter (linear phase filter)

C_lp=Cin.*fft(lp,n).';         %LP filter the response

Cout=C_lp+fft(hp,n).';          %add flat response to frequencies above 10
    kHz

end
```

B.3 General

B.3.1 createaudiofile.m

```
%% Binaural listening audio file creation, MSc-thesis Jonas Karlberg
    2016

% run CTCfilter_creation prior to this script

%% read audiofile

[x,Fs]=audioread('inputfile.wav');

x_l=x(:,1); %input signal left
x_r=x(:,2); %input signal right

%%

%%%%%%%%%%%%%%%% 2 channel system %%%%%%%%%%%%%%%%

%create the input signals
x_l_in=fftfilt(c_ll_44k_256,x_l)+fftfilt(c_rl_44k_256,x_r);
x_r_in=fftfilt(c_rr_44k_256,x_r)+fftfilt(c_lr_44k_256,x_l);

y=[x_l_in x_r_in];

%normalising
max_value = max(abs(y(:)));

y = y ./ max_value .* .99;

figure;
plot(real(y))

%% create wavfile, filename_filter_sppos_dummypos.wav
audiowrite('outputfile.wav',y,Fs);
```

B.3.2 setup.m

```
%% Setup file, add paths etc. MSc-thesis Jonas Karlberg 2016

%set folder structure
addpath('ctc_calc/smooth','ctc_calc','meas_data','measure','plots','
    evaluation','analysis')

%% play and record settings
info=audiodevinfo;          %fetch info from audio device(s)
IO=1;                       %select input
devices=audiodevinfo(IO);   %number of devices

if devices==1

    ID=info.input.ID;                       %sets device to internal
    name = audiodevinfo(IO,ID);
    IDin=0;                         %input
    IDout=1;                        %output

else if devices==2      %condition depends on soundcard

    ID=audiodevinfo(IO);                    %select device
```

```
    name = audiodevinfo(IO,ID);              %stores name of device
    name='default';
    IDin=ID;                                  %input device
    IDout=ID;                                 %output device


else
    ID=audiodevinfo(IO);                     %select device
    name = 'default';              %stores name of device

    IDin=3;                                  %input device
    IDout=3;                                 %output device
    end
end

disp(name)                   %prints the name of the used devicehelp aud
nChannels=2;          %1 mono, 2 stereo
bit=24;               %bitsample resolution
fs=48000;             %sample frequency

%% define variables
p_ref=20E-6;                                 %reference pressure
ref=1;                                       %reference value for filter
    plots

%% function calls
% here you can predefine variables for function calls
```

B.3.3 frfplots_general.m

```
%% FRF plots, MSc-thesis Jonas Karlberg 2016

function[]=frfplots_general(fs,p_ref,xmin,xmax,H1,H2,H3,H4)

n=length(H1);
f=linspace(0,fs,n+1);
f=f(1:end-1);

figure;
subplot(2,1,1)
semilogx(f,20.*log10(abs(H1)./p_ref),'g')
hold on, grid on
semilogx(f,20.*log10(abs(H4)./p_ref),'-.r')
axis([xmin xmax ((min(20.*log10(abs(H1)./p_ref))+min(20.*log10(abs(H4)
    ./p_ref)))/2-5) ((max(20.*log10(abs(H1)./p_ref))+max(20.*log10(abs(
    H4)./p_ref)))/2+5)])
legend('H_{1}','H_{4}','location', 'southwest')
title('Frequency response')
ylabel('magnitude'), xlabel('frequency')

subplot(2,1,2)
semilogx(f,20.*log10(abs(H2)./p_ref),'--b')
hold on, grid on
semilogx(f,20.*log10(abs(H3)./p_ref),'.-k')
axis([xmin xmax ((min(20.*log10(abs(H2)./p_ref))+min(20.*log10(abs(H3)
    ./p_ref)))/2-5) ((max(20.*log10(abs(H2)./p_ref))+max(20.*log10(abs(
```

```
    H3)./p_ref)))/2+5)])
legend('H_{2}','H_{3}','location', 'southwest')
title('Frequency response')
ylabel('magnitude'), xlabel('frequency')
end
```

B.3.4 frfplots2.m

```
%% FRF plots 2; plot function for comparing 2 TFs, MSc-thesis Jonas
    Karlberg 2016

function []=frfplots2(fs1,fs2,p_ref,xmin,xmax,H1,H2)

n1=length(H1);
n2=length(H2);
f1=linspace(0,fs1,n1+1);
f1=f1(1:end-1);
f2=linspace(0,fs2,n2+1);
f2=f2(1:end-1);

figure;
semilogx(f1,20.*log10(abs(H1)./p_ref),'g')
hold on, grid on
semilogx(f2,20.*log10(abs(H2)./p_ref),'-.r')
axis([xmin xmax ((min(20.*log10(abs(H1)./p_ref))+min(20.*log10(abs(H2)
    ./p_ref)))/2-5) ((max(20.*log10(abs(H1)./p_ref))+max(20.*log10(abs(
    H2)./p_ref)))/2+5)])
legend('H_1','H_2','location', 'southwest')
ylabel('magnitude'), xlabel('frequency')


end
```

B.4 Examples

The following code can be used as a starting point for the main script structure. Note that all functions used here are not included as part of the thesis. *CTC-filter_creation.m* can be used for creating filters with previously recorded data. *CTC_filters.m* includes structure for setup, measuring, filter creation etc. For analysis of captured data one may use *analysis.m*.

B.4.1 CTC_filters.m

```
%% CTC-filters; measure -> calculate -> plots, MSc-thesis Jonas
    Karlberg, 2016
clear all
close all
clc

%% initiation
setup
%%% FILING %%% (follwing to be used as name for saving data)
%%% enter date 'YYMMDD'
```

```
folderdate='161021';
%%% enter name of speaker system (e.g. '2ch_14carback', '2ch_carceil')
sysname='2ch_14carback';
%%% enter speaker position 'XX'
speakerpos='01';
%%% enter dummy position 'XX'
dummypos='01';

%%% save settings
savework=1;                                          %1=yes, 0=no, save workspace

%%% MEASURING %%%
%%% enter sweep length
sweeplength=2;                                    %sets length of sweep in s
%%% enter number of averages
average=4;                                        %sets the number of averages
%%% enter number of channels

%%% CALCULATION %%%
%%% calculate CTC-filters?
calc=0;                                          %1=yes, 0=no

%%% enter smoothing parameters
f_c=12000;      %fc filter "high all pass", 10 000, 20 000
ftrans=1000;    %sets the transistion frequency for smoothing
%%% frequency smoothing
N1=50;          %smoothing below ftrans, BACK 50, FRONT 100
N2=400;         %smoothing above ftrans, BACK 400, FRONT 200
%%% phase smoothing
N3=5;            %smoothing below ftrans, 100
N4=500;          %smoothing above ftrans, 4000

waudio=0;        %write audiofile? (filter IRs) 1=yes, 0=no

%%% apply allpass-filter or shelf-filter (with same gain as at point f_2)
allpass=2;       % 1=all-pass, 0=high shelf 2= LP-filter at 12 kHz

%%% enter length of IR
stop=2;                          %enter length of desired IR in s (for IR extraction)
%%% enter length of final IR (affects frequency response of CTC-filters)
cut=2^10;                              %sets the length for cutting IR in samples, to cut out reflections

%%% PLOTTING %%%
%%% enter freq range
xmin=100;                                        %define x-axis limits for FRF-plots
xmax=20000;                                      %define x-axis limits for FRF-plots
%%% enter time span
```

```
plotstop=4;                          %length of IR in ms (for IR plots axis), if 0 no limit on x-axis
%%% choose plots
plotIR=1;                            %1=yes, 0=no, plot IRs
plotTF=1;                            %1=yes, 0=no, plot TFs
plotCTC=1;                           %1=yes, 0=no, plot CTC TFs (if calc = 0, no plot)
plotCTCIR=1;                         %1=yes, 0=no, plot CTC IRs (if calc = 0, no plot)
plotCTCcomp=1;                       %1=yes, 0=no, plot CTC TF comparison (if calc = 0, no plot)
plotspec=1;                          %1=yes, 0=no, plot spectrograms of recorded material
%%%%%% NO EDITING BEYOND THIS LINE %%%%%%
%% filing
% create directories of input "folderdate"
mkdir('meas_data',folderdate)
mkdir('output/figures/',folderdate)
mkdir('output/filters/',folderdate)
mkdir('output/soundfiles/',folderdate)
mkdir('output/workspace/',folderdate)
mkdir('output/IRs/',folderdate)

irlength=num2str(cut); % storing irlength for filename

%%% filename format: xch_name_speakerpos_dummypos, ex: 2ch_back_01_01
rawdata=[sysname '_' dummypos '_' speakerpos '.mat'];

%%% stores the save path for raw data
savepath=['meas_data/' folderdate '/' rawdata];

%%% filename format: tfs_speakerpos_dummypos_IRlength, ex: tfs_01_01_231 (for an IR length of 231 samples)
tfs=['output/filters/' folderdate '/' sysname '_tfs_' speakerpos '_' dummypos '_' irlength '.mat']; %TFs of cut data

%%% filename format: CTC_filters_IRlength, ex: CTC_filters_231 (for an IR length of 231 samples)
filterpath=['output/filters/' folderdate '/' sysname '_CTC_filters_' irlength '.mat'];

workspace=['output/workspace/' folderdate '/' sysname '_CTC_filters.mat']; %enter name for storing complete workspace


%% measure

sweep=create_stimuli(fs,sweeplength);
disp('Prepare for measuring H_LL. Press any key to continue.')
pause
[H_LL,h_ll,audio_ll]=ir_capture(sweep,bit,fs,nChannels,IDin,IDout,stop,average); %measure H_LL

disp('Prepare for measuring H_LR. Press any key to continue.')
pause

```

```
[H_LR,h_lr,audio_lr]=ir_capture(sweep,bit,fs,nChannels,IDin,IDout,stop,
    average); %measure H_LR
disp('Prepare for measuring H_RL. Press any key to continue.')
pause

[H_RL,h_rl,audio_rl]=ir_capture(sweep,bit,fs,nChannels,IDin,IDout,stop,
    average); %measure H_RL
disp('Prepare for measuring H_RR. Press any key to continue.')
pause

[H_RR,h_rr,audio_rr]=ir_capture(sweep,bit,fs,nChannels,IDin,IDout,stop,
    average); %measure H_RR

%% store captured raw data
save(savepath, 'H_LL','H_LR','H_RL','H_RR','h_ll','h_lr','h_rl','h_rr',
    'fs','sweep','audio_ll','audio_lr','audio_rl','audio_rr');


%% fix IR of cross-talk paths
% perhaps add condition for peak? but will probably not occur before 60
% samples... but could use the same logic as in ir_capture using "peak"
h_lr=windctIR(h_lr);
h_rl=windctIR(h_rl);

%% cut and HP-filter IR before side door reflection

[H_LL,h_ll]=cutIR(fs,h_ll,cut);
[H_LR,h_lr]=cutIR(fs,h_lr,cut);
[H_RL,h_rl]=cutIR(fs,h_rl,cut);
[H_RR,h_rr]=cutIR(fs,h_rr,cut);

%% save cut TFs

save(tfs,'H_LL','H_LR','H_RL','H_RR','h_ll','h_lr','h_rl','h_rr');

%% calculate CTC-filters

if calc == 1

    [C_LL_ideal,C_LR_ideal,C_RL_ideal,C_RR_ideal]=ctc_calc_ideal(H_LL,
    H_LR,H_RL,H_RR);


    [C_LL,C_LR,C_RL,C_RR]=ctc_calc_smooth(N,H_LL,H_LR,H_RL,H_RR,fs,f_1,
    f_2);

    c_ll=(circshift(ifft(C_LL),[700 0]));          %IR of CTC filters
    c_lr=(circshift(ifft(C_LR),[700 0]));
    c_rl=(circshift(ifft(C_RL),[700 0]));
    c_rr=(circshift(ifft(C_RR),[700 0]));

    C_LL=fft(c_ll);
    C_LR=fft(c_lr);
    C_RL=fft(c_rl);
    C_RR=fft(c_rr);
else if calc == 0
```

```
    end
end

%% IR plots
if plotIR == 1
    irplots(sweep,fs,plotstop,h_ll,h_lr,h_rl,h_rr)
       %IR plots
else if plotIR == 0
    end
end

%% IR plots of CTC-filters
if plotCTCIR == 1 && calc == 1
    ctc_irplots(fs,plotstop,c_ll,c_lr,c_rl,c_rr)
else if plotCTCIR == 0
    end
end

%% FRFplots of propagation paths and CTC-filters

if plotTF == 1
    frfplots(fs,p_ref,xmin,xmax,H_LL,H_LR,H_RL,H_RR)                    %
    frequency response plots of propagation paths
else if plotTF == 0
    end
end

if plotCTC == 1 && calc == 1
    ctcplots(fs,ref,xmin,xmax,C_LL,C_LR,C_RL,C_RR)                    %
    frequency response plots of calculated CTC filters
else if plotCTC == 0
    end
end

%% comparison IDEAL and smoothed filters
if plotCTCcomp == 1 && calc == 1
    ctc_comp(fs,ref,xmin,xmax,C_LL,C_LR,C_RL,C_RR,C_LL_ideal,C_LR_ideal
    ,C_RL_ideal,C_RR_ideal)
else if plotCTCcomp == 0
    end
end
%% plot spectograms
if plotspec == 1
    figure; spectrogram(audio_ll(:,1),512,256,512,48000,'yaxis')
    title('Direct path Left')
    figure; spectrogram(audio_lr(:,1),512,256,512,48000,'yaxis')
    title('Cross talk, Left -> Right')
    figure; spectrogram(audio_rl(:,1),512,256,512,48000,'yaxis')
    title('Cross talk, Right -> Left')
    figure; spectrogram(audio_rr(:,1),512,256,512,48000,'yaxis')
    title('Direct path Right')
else if plotspec == 0
    end
end

%% save workspace
```

```
if savework == 1
    save(workspace)
else if savework == 0
    end
end

%% save filters
if calc == 1
    save(filterpath ,'C_LL','C_LR','C_RL','C_RR','N')
else if calc == 0
    end
end
```

B.4.2 CTCfilter_creation.m

```
%% CTC-filter creation, MSc-thesis Jonas Karlberg, 2016
% load the TFs that the filters should be based on

fs=48000;    %sample frequency

f_c=12000;      %fc filter "high all pass", 10 000, 20 000
ftrans=1000;    %sets the transistion frequency for smoothing
%%% frequency smoothing
N1=500;            %smoothing below ftrans, 50
N2=1000;           %smoothing above ftrans, 400
%%% phase smoothing
N3=5;              %smoothing below ftrans, 100
N4=500;            %smoothing above ftrans, 4000

waudio=0;          %write audiofile? (filter IRs) 1=yes, 0=no

%%% apply allpass-filter or shelf-filter (with same gain as at point
    f_2)
allpass=2;         % 1=all-pass, 0=high shelf 2= LP-filter at 12 kHz

%%% for function call
FRFs=[H_LL_cut,H_LR_cut,H_RL_cut,H_RR_cut]; % measured frequency
    responses
smoothing=[N1,N2,N3,N4];                            %smoothing parameters

%%% for plotting
xmin=100;
xmax=20000;
p_ref=1;

%% Calculate filters
[C_LL_short,C_LR_short,C_RL_short,C_RR_short]=ctc_calc_smooth5(
    smoothing,ftrans,FRFs,fs,f_c,allpass);
%% create IRs
c_ll_short=real(ifft(C_LL_short));
c_lr_short=real(ifft(C_LR_short));
c_rl_short=real(ifft(C_RL_short));
c_rr_short=real(ifft(C_RR_short));

n=length(c_ll_short);

```

```
c_ll_short_flip=(circshift(real(ifft(C_LL_short)),[round(n/2) 0]));
        %IR of CTC filters
c_lr_short_flip=(circshift(real(ifft(C_LR_short)),[round(n/2) 0]));
c_rl_short_flip=(circshift(real(ifft(C_RL_short)),[round(n/2) 0]));
c_rr_short_flip=(circshift(real(ifft(C_RR_short)),[round(n/2) 0]));


%%% apply window

fadein= hann(2^10);
fadeout= hann(2^10);

fade=[fadein( 1:end/2 ); fadeout(end/2+1:end)];

window = [ fadein( 1:end/2 ); ones( n - length(fade), 1 ); fadeout(end
    /2+1:end) ];

c_ll_short_flip=c_ll_short_flip.*window;
c_lr_short_flip=c_lr_short_flip.*window;
c_rl_short_flip=c_rl_short_flip.*window;
c_rr_short_flip=c_rr_short_flip.*window;

%% create audiofiles

%%% downsample to 44.1 kHz (for use in SSR)
[P,Q] = rat(44100/fs);
abs(P/Q*fs -44100);

c_ll_44k_256 = resample(c_ll_short_flip,P,Q);
c_lr_44k_256 = resample(c_lr_short_flip,P,Q);
c_rl_44k_256 = resample(c_rl_short_flip,P,Q);
c_rr_44k_256 = resample(c_rr_short_flip,P,Q);

%%%

filter_1 = c_ll_44k_256;      %C_LL
filter_2 = c_lr_44k_256;      %C_LR

filter_3 = c_rl_44k_256;      %C_RL
filter_4 = c_rr_44k_256;      %C_RR

%%% normalize IRs, uses the greatest value of max(abs(filter_x)))

if max(abs(filter_1)) > max(abs(filter_4))
    norm=max(abs(filter_1))
else
        norm=max(abs(filter_4))
end

filter_1norm=filter_1./(max(abs(norm))*1.01);
filter_2norm=filter_2./(max(abs(norm))*1.01);

filter_3norm=filter_3./(max(abs(norm))*1.01);
filter_4norm=filter_4./(max(abs(norm))*1.01);

if waudio == 1
```

```

    audiowrite( 'output/IRs/impulse_responses_in1_to_out1_and_in1_to_out2.wav', [ filter_1norm, filter_2norm ], 44100, 'BitsPerSample', 24 );
    audiowrite( 'output/IRs/impulse_responses_in2_to_out1_and_in2_to_out2.wav', [ filter_3norm, filter_4norm ], 44100, 'BitsPerSample', 24 );

end

frfplots_general(44100,p_ref,xmin,xmax,fft(filter_1norm,2^18),fft(filter_2norm,2^18),fft(filter_3norm,2^18),fft(filter_4norm,2^18))
```

B.4.3 analysis.m

```
%% Analysis; view captured data, MSc-thesis Jonas Karlberg, 2016
clear all
close all
clc

%% setup %%

%%% FILING %%% (follwing to be used for loading RAW data)
%%% enter date 'YYMMDD'
folderdate='161021';
%%% enter name of speaker system (e.g. '2ch_14carback', '2ch_carceil')
sysname='2ch_14carback';
%%% enter speaker position 'XX'
speakerpos='01';
%%% enter dummy position 'XX'
dummypos='01';


%%% IR EXTRACTION %%%
stop=2;                                 %set IR length in s
fs=48000;                               %set sample rate

%%% for cutting loaded IR %%%
%%% enter filter cut length
cut=2^9;        %back system 2^8, front 2^7

%%% PLOTTING %%%

%%% enter freq range
xmin=100;                               %define x-axis limits for FRF-plots
xmax=20000;                             %define x-axis limits for FRF-plots
p_ref=1;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

irlength=num2str(cut);  % storing irlength for filename

%%% filename format: xch_name_speakerpos_dummypos, ex: 2ch_back_01_01
rawdata=[sysname '_' dummypos '_' speakerpos '.mat'];

```

```
%%% stores the save path for raw data
savepath=['meas_data/' folderdate '/' rawdata];

%% load raw data

load(savepath);

%% fetch new transfer function with different cut length

[~,h_ll_long]=ir_extraction(audio_ll,fs,stop);     %HP filter at 100 Hz
[~,h_lr_long]=ir_extraction(audio_lr,fs,stop);
[~,h_rl_long]=ir_extraction(audio_rl,fs,stop);
[~,h_rr_long]=ir_extraction(audio_rr,fs,stop);

[h_ll_cut]=cutIR2(fs,h_ll_long,cut);        %cut IR
[h_lr_cut]=cutIR2(fs,h_lr_long,cut);
[h_rl_cut]=cutIR2(fs,h_rl_long,cut);
[h_rr_cut]=cutIR2(fs,h_rr_long,cut);


%% add zeros to IR

h_zeros=zeros((2^12),1);     %determines length of filter

h_ll_adz=[h_ll_cut; h_zeros(1:end-cut,1)]; %adding zeros to IR
h_lr_adz=[h_lr_cut; h_zeros(1:end-cut,1)];
h_rl_adz=[h_rl_cut; h_zeros(1:end-cut,1)];
h_rr_adz=[h_rr_cut; h_zeros(1:end-cut,1)];

%% compute FFTs

H_LL_long_hp=fft(h_ll_long);
H_LR_long_hp=fft(h_lr_long);
H_RL_long_hp=fft(h_rl_long);
H_RR_long_hp=fft(h_rr_long);

H_LL_cut=fft(h_ll_adz);
H_LR_cut=fft(h_lr_adz);
H_RL_cut=fft(h_rl_adz);
H_RR_cut=fft(h_rr_adz);

%% plot FRFs

%%% FRFs for analysis
frfplots_general(fs,p_ref,xmin,xmax,H_LL_long_hp,H_LR_long_hp,
    H_RL_long_hp,H_RR_long_hp)
frfplots_general(fs,p_ref,xmin,xmax,H_LL_cut,H_LR_cut,H_RL_cut,H_RR_cut
    )
```