M.Sc. Thesis
Master of Science in Engineering Acoustics

Sensitivity of the sound zones problem to sources of error

Colin Benker (s190247)

Kongens Lyngby, February 1, 2021

DTU Acoustic Technology
Department of Electrical Engineering
Technical University of Denmark

Elektrovej
Building 352
2800 Kongens Lyngby, Denmark
Phone (+45) 4525 3949
www.act.elektro.dtu.dk

Abstract

Sound zones systems generate a sound field in a spatially confined zone while attenuating the sound energy in another predefined zone within the same acoustic space. The goal of these systems is to enable different listeners to enjoy their individual audio content without disrupting each other or having to wear headphones.

It was shown in the literature, that the performance of sound zones is sensitive to changes in the physical environment, such as changes in the speed of sound due to temperature changes [OM17] or loudspeaker transfer function changes [Ma+18; Ma+19; PCK13]. However, current research papers do not provide an estimation of how severe the influence of transfer function errors is on the sound zones problem at low frequencies in domestic rooms. This thesis provides an error model, that simulates transfer function errors based on physical properties of a room and loudspeakers. This error model is incorporated in a Monte-Carlo simulation to provide an evaluation of the effect of transfer function changes on the acoustic separation between the zones. This effect is evaluated in terms of the acoustic contrast, an energy ratio between the two sound zones [CK02]. The frequency range covered in this thesis is the low frequency range from 20 to 300 Hz.

The results from the simulations suggest that changes in room temperature have a significant impact on the acoustic contrast in the whole frequency range, whereas the impact of loudspeaker transfer function changes is most prominent at very low frequencies below 100 Hz. Above 150 Hz the loudspeaker errors only decreased the performance marginally.

Further, this thesis investigated how the sound zones system can be made more robust to transfer function changes. A main conclusion from the conducted simulations is, that the more information about the distributions of the underlying transfer function changes is accessible, the better the system can be regularized.

Preface

This Master's thesis was prepared at the Acoustic Technology group in the department of Electrical Engineering at the Technical University of Denmark in fulfillment of the requirements for acquiring a M.Sc. degree in Engineering Acoustics.

Kongens Lyngby, February 1, 2021

Colin Benker (s190247)

Acknowledgements

I would like to thank Martin Bo Møller and Efren Fernandez-Grande for the amazing supervision of this thesis. The frequent discussions we had inspired me immensely and had an incredibly valuable impact on the course of this project. I am very grateful for all the time they dedicated to giving me scientific insights and assisting me in shaping this thesis.

List of acronyms

AC Acoustic Contrast

ACC Acoustic Contrast Control

MC Monte-Carlo

OLS Ordinary Least Squares

PM Pressure Matching

PMO Probability Model Optimization

RTLS Regularized Total Least Squares

RR Ridge Regression

SPL Sound Pressure Level

SVD Singular Value Decomposition

TF Transfer Function

TLS Total Least Squares

List of symbols

In general, scalars are represented by italic lower- or uppercase letters (e.g. a or A), one-dimensional vectors are represented by lowercase bold letters (e.g $\boldsymbol{a}$), and higher order matrices are denoted by upper case bold letters (e.g. $\boldsymbol{A}$).

Signal Path

$h_{l,s}$	Loudspeaker transfer function of loudspeaker l
$h_{ml,r}$	Room transfer function
l	Index of loudspeaker
L	Number of loudspeakers
m	Index of microphone
M	Number of microphones
p_m	Sound pressure at position of microphone m
q_l	Volume velocity emitted by loudspeaker l
U_{out}	Output voltage of Amplifier
w_l	Sound Zones Filter of loudspeaker l
x	Input Signal

Loudspeaker Transfer Function

Bl	Force factor of coil-magnet assembly
C_{AB}	Air-load compliance in the box
C_{AS}	Compliance of suspension
C_{AT}	Total acoustic compliance
i_c	Current flowing through voice coil
M_{AB}	Acoustic mass of air moving inside box
M_{AF}	Acoustic mass of air oscillating on the front of diaphragm
M_{AS}	Acoustic mass of moving coil and diaphragm
M_{AT}	Total moving acoustic mass
M_{AT}	Total acoustic mass moving
p_D	Pressure difference inside and outside box
Q_{TC}	Total closed box quality factor

R_{AS}	Resistance modelling mechanical losses in diaphragm
R_{AT}	Total acoustic resistance
R_E	Initial DC-Resistance of coil
$\tilde{R_E}$	Temperature dependent DC-Resistance of coil
s	Complex frequency
S_D	Surface area of diaphragm
U_{coil}	Voltage induced in coil
v_D	Velocity of diaphragm
ω_c	Total closed box angular resonance frequency
Z_{AB}	Acoustical impedance inside the box
Z_{AF}	Acoustical impedance at the front of the box
Z_{AT}	Sum of acoustic impedances
Z_E	Electrical impedance
Z_M	Mechanical impedance of system

Room Transfer Function

c	Speed of sound
f	Frequency
G_{ml}	Green's function between microphone m and loudspeaker l
j	Imaginary unit, $\sqrt{-1}$
k	Wave number
k_ν	Natural wavenumber of mode ν
$\boldsymbol{r}_l$	Position of loudspeaker l
$\boldsymbol{r}_m$	Position of microphone m
$\delta(\cdot)$	Dirac delta function
ψ_ν	Mode shape of mode ν
ρ_0	Density of air
τ_ν	Time constant of mode ν
ω	Angular frequency

Sound Zones Problem

$\boldsymbol{p}$	Vector of sound pressures p_m
$\boldsymbol{H}_r$	Room transfer function matrix
$\boldsymbol{h}_s$	Vector of speaker transfer functions $h_{l,s}$
$\boldsymbol{w}$	Vector of sound zones filters w_l
$\boldsymbol{H}$	Total transfer function matrix
$\boldsymbol{I}$	Identity matrix

Least Squares Problem

$\boldsymbol{A}$	Data matrix
$\boldsymbol{a}_1, \boldsymbol{a}_2, ..$	Column vectors of $\boldsymbol{A}$
$\boldsymbol{b}$	Vector of observations
$\hat{\boldsymbol{b}}$	Projection of $\boldsymbol{b}$ on span of $\boldsymbol{A}$
r	Rank of matrix
$\boldsymbol{r}$	Residual vector $\boldsymbol{b} - \hat{\boldsymbol{b}}$
$\boldsymbol{U}_i$	Left singular matrix
$\boldsymbol{u}_i$	Left singular vector of singular value decomposition
$\bar{\boldsymbol{u}}_i$	Left singular vector of $[\boldsymbol{A}, \boldsymbol{b}]$
$\boldsymbol{V}_i$	Right singular matrix
$\boldsymbol{v}_i$	Right singular vector of singular value decomposition
$\bar{\boldsymbol{v}}_i$	Right singular vector of $[\boldsymbol{A}, \boldsymbol{b}]$
σ_i	Singular values
$\kappa(\cdot)$	Condition number of matrix
λ	Regularization parameter
$\bar{\sigma}_i$	Singular value of $[\boldsymbol{A}, \boldsymbol{b}]$
$\boldsymbol{x}$	Weight vector

Simulations

$\boldsymbol{H}_0$	Initial Transfer function matrix
$\boldsymbol{H}_R$	Set of transfer function matrices with room transfer function errors
$\boldsymbol{H}_S$	Set of transfer function matrices with speaker transfer function errors
$\boldsymbol{H}_{SR}$	Set of transfer function matrices with speaker and room transfer function errors
$\boldsymbol{w}_{0,res,err}$	Resampled filters based on $\boldsymbol{w}_0$ and flawed temperature measurements
$\boldsymbol{w}_{0,res}$	Resampled filters based on $\boldsymbol{w}_0$
$\boldsymbol{w}_{PMO,R}$	PMO filters based on $\boldsymbol{H}_R$
$\boldsymbol{w}_{PMO,S}$	PMO filters based on $\boldsymbol{H}_S$
$\boldsymbol{w}_{PMO,SR}$	PMO filters based on $\boldsymbol{H}_{SR}$
$\boldsymbol{w}_{PMO_{ME}}$	PMO filters assuming multiplicative errors
$\boldsymbol{w}_{PMO_{Over}}$	PMO filters based on overestimated transfer function errors
$\boldsymbol{w}_{PMO_{Under}}$	PMO filters based on overestimated transfer function errors
$\boldsymbol{w}_{RR,res}$	Resampled filters based on $\boldsymbol{w}_{RR}$
$\boldsymbol{w}_{RR}$	Ridge regression filters
$\boldsymbol{w}_{TLS}$	TLS filters
$\boldsymbol{w}_0$	Initial PM filters, based on OLS

Contents

CHAPTER 1

Introduction

This chapter provides the reader with an introduction to the sound zones problem in general, and specifically to the research questions addressed in this thesis. The first section of this chapter, the motivation, introduces the sound zones problem and describes *why* this research was carried out. The second section, the scope of the project, introduces the addressed research questions and describes *what* this project contributes to current research. Section 1.3 presents the outline of this thesis, i.e. *how* this thesis was structured to discuss the presented research questions.

1.1 Motivation

So far, whenever multiple individuals within the same acoustic space wish to enjoy independent audio content, they have to rely on headphones to prevent interfering audio from distracting each other. Headphones however, impede communication between listeners within that space. It is therefore desired to create a loudspeaker system that delivers individual audio content to each listener, without distracting interference [Col+14b]. This can be achieved by creating so-called *sound zones* within the room. The idea is to create a bright and a dark zone for each individual audio content. The bright zone thereby denotes a region within the space, where the audio content is played back, while it is suppressed as much as possible in the dark zone. Outside these two zones the sound field is not controlled. To create two zones, where two individual audio contents are played back at the same time, one can superimpose two sound zone solutions. For example, when two listeners sit in a region A and a region B in a room, two sound zone solutions need to be created. One that generates a dark zone in region A and a bright zone in region B, and one solution that creates a dark zone in region B and a bright zone in region A.

Possible use case scenarios of the sound zones are for example: different people watching the same movie in different languages at the same time, someone reading a book in silence while someone else is watching TV, or different people listening to different music at the same time.

One way to create sound zones is by using a composite solution, which uses active sound field control in the low frequency range, array processing at mid frequencies and the directivity of tweeters at high frequencies, as introduced by Druyvesteyn et al. in [DG97]. This thesis focuses on the low frequency approach in the range from 20-300 Hz. In this range, where the wavelengths are long, sound zones can be achieved through interference. To obtain silence in the dark zone, the sound from a number of loudspeakers has to interfere destructively in the dark zone. In Pressure Matching

(PM), the sound in the bright zone is desired to interfere in such a way that the sound field equals that of a predefined target sound field [KN93]. Pressure Matching is a sound zones method that aims at minimizing the mean square error between a desired target sound field and the actually reproduced sound field. To achieve the desired interference pattern in the room, and thereby the sound zones, the phase and magnitude of the input signal have to be controlled for each loudspeaker. This can be accomplished by applying pressure matching filters to the input signal.

In order to design pressure matching filters, the transfer functions from all sound sources to the targeted position need to be known. These transfer functions contain the information about how the system, i.e. the loudspeakers and the room, alter the input signal phase and magnitude, before the signal can be perceived or recorded as a sound pressure in the sound zones. The transfer functions from each loudspeaker to any point in a real room are unique. A given input signal will therefore arrive with a different phase and magnitude at a given position, depending on the location and impulse response of the speaker emitting the signal. When all the individual transfer functions are known, pressure matching filters can find the linear combination of these transfer functions, that minimizes the squared error between the actually reproduced sound field and the desired sound field.

Sound zones systems in rooms are usually realized as feed-forward systems based on measured transfer functions [Møl+19]. This poses two challenges to the design of PM filters. First, the transfer function measurements will always contain noise, such that the transfer functions used for the filter calculation are only an estimate of the real underlying transfer functions. Møller et al. [Møl+19] showed, that this measurement noise can significantly degrade the performance of sound zones systems. Secondly, the transfer functions are not static. Transfer functions can change after they were determined, as a result to changes in the loudspeaker or in the environment. These transfer function errors can have a significant impact on the sound zones performance. The sensitivity of the performance to transfer function changes was investigated in various setups. Olsen et al. [OM17] for example investigated the effect of transfer function errors due to changes in the ambient temperature in cars. Chang et al. [CJ12] investigated which effects scattering, e.g. by a human head, can have on the performance of sound zones systems. The effect of phase and magnitude errors in the loudspeakers, and the effect of small position mismatches is described by Park et al. in [PCK13]. Further Ma et al. [Ma+18; Ma+19] investigated the impact of loudspeaker distortion on the sound zones problem.

While these studies gave insights to how transfer function changes can impact sound zones performance, it is difficult to draw conclusions on how severe the effects are when sound zones are realized in domestic rooms. The mentioned study by Olsen et al. [OM17] for example investigated the effects of a temperature change from -2 °C to 22.8 °C. In domestic rooms the temperature variations can be assumed to be much smaller. Further, the interior of a car has different properties than domestic rooms in terms of size and damping. Ma et al. [Ma+19] showed that deviations between estimated and real loudspeaker transfer functions can have a significant impact on the sound zones performance, but since this paper focuses on the effects of distortion, it was not investigated how severe the impact of loudspeaker TF errors is across frequency. The sensitivity to phase and magnitude errors in loudspeakers is described analytically in Park et al. [PCK13], but since the setup and sound zones method are different these results can not be directly compared to domestic rooms.

To increase the robustness of the sound zones problem to various sources of error, regularization strategies were introduced. Zhu et al. [Zhu+17] provides a comparison of different regularization strategies on their ability to make the system robust to additive or multiplicative noise. The noise parameters are designed arbitrarily, i.e. the noise is not based on a physical model. A physical error model however, could provide further insights to the effectiveness of regularization, when the sound zones are realized in domestic rooms.

1.2 Scope of the project

This thesis aims at providing realistic estimations of the challenges in the low frequency sound zones problem in domestic rooms. In order to acquire these insights an error model is proposed, that models transfer function errors. Two main error sources are considered here: The change of room temperature altering the room transfer functions, and loudspeaker transfer function changes as a result of the voice coil heating up. These error models are then incorporated in a Monte-Carlo (MC) simulation to investigate two research questions. The first research question asks:

"Which impact do transfer function errors have on the performance of a low frequency sound zones system?"

This impact is evaluated by means of the acoustic contrast, an energy ratio between the bright and the dark zone [CK02]. It is investigated how the acoustic contrast is affected by the discrepancy between the transfer function estimates the PM filters are based on, and the actual transfer functions they are evaluated on. This discrepancy arises due to the actual transfer functions changing in each MC iteration, which is simulated by the error model.

The second research question is concerned with regularization. It is of interest to investigate how the sound zones system can be made more robust to transfer function errors. The second research question asks:

"How can different regularization strategies improve the robustness of a low frequency sound zones system to transfer function errors?"

The different regularization strategies aim at reducing the sensitivity to transfer function errors, by refining the PM filters. The different regularization strategies covered here are ridge regression (RR) [Zhu+17; Col+14b; The15], total least squares (TLS) [HO96; The15; GL13], resampling [OM17], and probability model optimization (PMO) [Zhu+17]. Again, the performance of the different regularization strategies is compared by means of the acoustic contrast in a MC simulation based on the provided error model.

By investigating these two research questions this thesis makes another step to analyzing the realisability of sound zones in domestic room settings.

1.3 Outline

Chapter 2 of this thesis introduces the theoretical background required to argue about the simulation introduced in chapter 3. As stated before, this thesis introduces an error model that allows an estimation on the impact of transfer function errors in the sound zones problem. Therefore, the first part of the theory section is concerned with providing the reader insights into the modelling of loudspeaker and room transfer functions. Thereafter, the sound zones problem and common challenges to the sound zones problem are introduced. In mathematical terms, pressure matching is a least squares (LS) problem. Therefore some basic properties of least squares problems are introduced in section 2.4, before the regularization strategies leveraged in the following simulations are introduced in section 2.5.

Chapter 3 explains how the transfer function errors are modeled in the MC simulation. Further it introduces how the different regularization strategies are implemented.

The results from the MC simulation are presented in chapter 4. First, the influence of transfer function errors on the ordinary PM filters is investigated, before the performance of different regularized PM filters is shown. Table 4.1 provides a summary of the acoustic contrast results of the different filters.

Chapter 5 discusses the results and limitations of the performed simulations and addresses how the stated research questions were accounted for.

Finally this thesis is concluded in chapter 6, by a brief summary of the results, some concluding remarks and an outlook to possible research that could be conducted in the future.

CHAPTER 2

Theoretical Background

The following chapter presents the necessary background theory to understand current challenges in the sound zones problem. The first section introduces the signal path, i.e. how the electrical signal is transformed to a volume velocity based on a model by Leach et al. [LL03] and how this volume velocity relates to the sound pressure at a desired position in a room. This knowledge is then applied to the context of sound zones in section 2.2, before some of the challenges in the sound zones problem are introduced in section 2.3. In this thesis Pressure Matching is used to realize sound zones in rooms. The PM solution is obtained by solving a least squares (LS) problem, so the concept of LS is presented in section 2.4 followed by common regularization strategies.

2.1 Signal Path

When the sound zones problem is implemented in situ, the transfer functions are determined from measurements. Thereby it has to be considered that the room transfer functions are only one part of the system that relates the input signal x to the recorded sound pressure p_m. In the first step of the signal path, the signal x is filtered by a (Pressure Matching) filter w_l. The output signal is scaled by an amplifier to the output voltage U_{out}. A loudspeaker l then transforms the output voltage of the amplifier U_{out} with the speaker transfer function $h_{l,s}$ to a volume velocity q_l. This volume velocity is then related to the sound pressure p_m at a microphone m by the room transfer function $h_{ml,r}$. The sound pressure is then transformed to an input voltage by a microphone which is recorded by the sound card. Each of these steps has a unique transfer function, however the transfer functions of the soundcard, amplifier and microphones are assumed to be flat and robust. Consequently the focus of this thesis lies on the transfer functions that relate the output voltage of the amplifier to the volume velocity, i.e. the speaker transfer function $h_{l,s}$, and the room transfer function relating the volume velocity emitted by the speaker to the sound pressure at the microphone $h_{ml,r}$. A block diagram summarizing the signal path is illustrated in figure 2.1.

All transfer functions are assumed to be linear and time-invariant. This thesis analyses the sound zones problem in the frequency domain, so the total transfer function from the output signal x to the recorded sound pressure p_m is a product of the individual transfer functions [PM06]. Therefore the individual transfer functions of the loudspeaker, room and filter can be analysed independently.

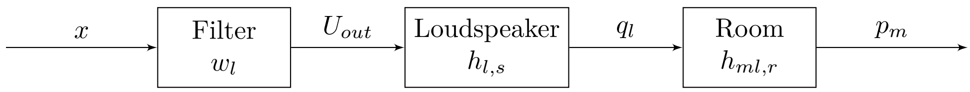


Figure 2.1: Signal Path in the sound zones system. An audio signal x is filtered by the filter w_l, the speaker transfer function $h_{l,s}$ and the room transfer function $h_{ml,r}$, resulting in a sound pressure p_m. The sound pressure p_m is transduced by microphones and recorded by a soundcard. The transfer functions of the amplifier, the microphones and the soundcard are assumed to be flat and are therefore not illustrated.

2.1.1 Loudspeaker transfer function

The transfer function of a loudspeaker relates the output voltage of an amplifier U_{out} to the volume velocity q_l generated by the movement v_D of the diaphragm with surface area S_D. In the following, the low frequency transfer function $h_{l,s}$ of a closed-box loudspeaker is derived.

The diaphragm of a speaker is attached to a coil in a magnetic field generated by a permanent magnet. The movement of the diaphragm is a result of the force acting on the coil when a current i_c flows through it [LL03]. This current can be expressed by [LL03]

$$i_c = \frac{U_{out} - U_{coil}}{Z_E} = \frac{U_{out} - Blv_D}{Z_E} = \frac{U_{out} - Bl\frac{q_l}{S_D}}{Z_E}. \tag{2.1}$$

U_{coil} denotes the induced voltage across the coil as a result of the movement of the coil with velocity v_D. In this simplified model, the electrical impedance Z_E can be approximated to the DC-Resistance of the coil R_E [LL03] at low frequencies. Bl denotes the force factor of the coil-magnet assembly which relates the mechanical force on the coil to the current flowing through it. The resulting diaphragm velocity v_D can be related to the mechanical forces generated by the current flowing through the coil and the pressure differences p_D inside and outside the box in terms of the mechanical impedance Z_M of the system [LL03]

$$v_D = \frac{Bli_c - S_D p_D}{Z_M} = \frac{Bli_c - S_D^2 v_D(Z_{AF} + Z_{AB})}{Z_M}. \tag{2.2}$$

The pressure difference was substituted using the acoustical impedances at the front of the box Z_{AF} and inside the box Z_{AB}. Inserting equation 2.1 into equation 2.2 and solving for the volume velocity yields [LL03]

$$q_l = v_D S_D = \frac{Bl}{R_E S_D} U_{out} \Big(\frac{Z_M}{S_D^2} + Z_{AF} + Z_{AB} + \frac{(Bl)^2}{S_D^2 R_E} \Big)^{-1}. \tag{2.3}$$

The first part of this equation can be seen as the pressure on the diaphragm resulting from the voltage on the coil. The second, inverted term is a sum of acoustic impedances relating the pressure to the generated flow, i.e. the volume velocity, of the system. This sum of acoustic impedances can be written as [LL03]

$$Z_{AT} = M_{AT}s + R_{AT} + \frac{1}{C_{AT}s}, \tag{2.4}$$

where M_{AT} is the total moving acoustic mass, R_{AT} is the total acoustic resistance, C_{AT} is the total acoustic compliance and s the complex frequency. M_{AT} is the sum of the effective acoustic mass M_{AB} of the air moving inside the box, the acoustic mass of the moving coil and diaphragm M_{AS} and the acoustic mass of the air oscillating on the front of the diaphragm M_{AF}. So M_{AT} is the total acoustic mass moving when a signal is passed to the speaker.

At low frequencies two approximations can be made. First, one can neglect the acoustic resistance and compliance of the air elements outside the box, since their impedances are much greater than the impedance of the air mass in parallel to them [LL03]. Second, the impedance in series caused by the coil inductance is small and can also be neglected at low frequencies [LL03]. This low frequency approximation holds for frequencies that are less than one half of the upper piston frequency limit [LL03].

The total acoustic resistance R_{AT} is simply the sum of the resistance modelling the mechanical losses in the diaphragm and coil suspension R_{AS}, the mechanical losses from the damping inside the box R_{AB} and the electrical losses induced by the coil movement in a magnetic field R_{AE}.

The total acoustic compliance C_{AT} is composed of the compliance of the suspension C_{AS} and the air-load compliance in the box C_{AB}.

The individual summands of the total acoustic impedance of the speaker can therefore be expressed as [LL03]

$$M_{AT} = M_{AB} + M_{AS} + M_{AF} \tag{2.5}$$

$$R_{AT} = R_{AS} + R_{AB} + R_{AE} \tag{2.6}$$

$$C_{AT} = \frac{C_{AS}C_{AB}}{C_{AS} + C_{AB}} \tag{2.7}$$

Inserting the acoustic impedances into equation 2.3 and reformulating the equation yields

$$\begin{aligned} q_l &= \frac{Bl}{R_E S_D \frac{1}{M_{AT}s + R_{AT} + \frac{1}{C_{AT}s}}} U_{out} \\ &= \frac{Bl}{R_E S_D} \frac{C_{AT}s}{M_{AT}C_{AT}s^2 + R_{AT}C_{AT}s + 1} U_{out} \\ &= \frac{Bl}{R_E S_D} \frac{1}{R_{AT}} \frac{\frac{1}{Q_{TC}}\frac{s}{\omega_c}}{(\frac{s}{\omega_c})^2 + \frac{1}{Q_{TC}}\frac{s}{\omega_c} + 1} U_{out} \end{aligned} \tag{2.8}$$

Where the total closed box quality factor Q_{TC} and the total closed box angular resonance frequency w_c were used to simplify the expression. They are defined as

[LL03]

$$Q_{TC} = \frac{1}{R_{AT}}\sqrt{\frac{M_{AT}}{C_{AT}}} \tag{2.9}$$

and

$$\omega_c = \frac{1}{\sqrt{M_{AT}C_{AT}}}. \tag{2.10}$$

Finally the volume velocity in equation 2.8 can be rewritten as

$$q_l = h_{l,s}U_{out}, \tag{2.11}$$

where the speaker transfer function is defined as

$$h_{l,s} = \frac{Bl}{R_E S_D}\frac{1}{R_{AT}}\frac{\frac{1}{Q_{TC}}\frac{s}{\omega_c}}{(\frac{s}{\omega_c})^2 + \frac{1}{Q_{TC}}\frac{s}{\omega_c} + 1}. \tag{2.12}$$

This low frequency transfer function can be used to relate the radiated volume velocity q_l to the output voltage U_{out} when the required parameters are known.

2.1.2 Room transfer function

The relation between the sound pressure p_m at a microphone position $\boldsymbol{r}_m$ and the volume velocity q_l emitted by a loudspeaker at a position $\boldsymbol{r}_l$ can be expressed by a room transfer function $h_{ml,r}$ as

$$p_m = h_{ml,r}q_l. \tag{2.13}$$

Room transfer functions can either be determined in situ from room impulse response measurements [Gui+15] or simulated. This report focuses on the simulation. To determine the transfer function $h_{ml,r}$ analytically one makes use of three fundamental acoustic principles: mass is conserved, harmonic pressure differences are compensated by changes in inertia and sound can be seen as an adiabatic phenomenon [JJ13]. These principles can be used to formulate the inhomogeneous Helmholtz equation

$$\nabla^2 p_m + k^2 p_m = -j\omega\rho_0 q_l \delta(\boldsymbol{r}_m - \boldsymbol{r}_l), \tag{2.14}$$

where $\omega = 2\pi f$ is the angular frequency, k the wave number $\frac{\omega}{c}$, c the speed of sound, ρ_0 the density of air and $\delta(\boldsymbol{r}_m - \boldsymbol{r}_l)$ is the Dirac delta function which is one for $\boldsymbol{r}_m = \boldsymbol{r}_l$ and zero otherwise. Reformulating the inhomogeneous Helmholtz equation by using Green's function G_{ml}, such that $p_m = j\omega\rho_0 q_l G$ results in

$$\nabla^2 G_{ml} + k^2 G_{ml} = -\delta(\boldsymbol{r}_m - \boldsymbol{r}_l). \tag{2.15}$$

Green's function is the general solution to the inhomogeneous Helmholtz equation and can be understood as the room response of a surrounding space. Depending on the boundary conditions of the space one can formulate specific Green's functions, e.g.

for free-field conditions, semi-infinite ducts or rooms [JJ13]. In rooms the boundary conditions are defined by the damping properties of the walls. In a lightly dampened room Green's function can be expressed as a sum of multiple room modes ν with a distinct mode shape ψ_ν and natural wavenumber k_ν

$$G_{ml}(f) = G(\boldsymbol{r}_m, \boldsymbol{r}_l, f) = \frac{-1}{V} \sum_{\nu=0}^{\infty} \frac{\psi_\nu(\boldsymbol{r}_m)\psi_\nu(\boldsymbol{r}_l)}{k^2 - k_\nu^2 - jk/(\tau_\nu c)}. \tag{2.16}$$

The imaginary term $-jk/(\tau_\nu c)$ expresses the damping of the mode, where τ_ν is the time constant of the ν^{th} mode, which is proportional to the reverberation time at the corresponding natural frequency. In the simplified example of a perfectly rectangular room without damping the mode shapes ψ_ν are simply a scaled product of cosines in the x, y and z dimensions with maximums at the walls. Mode shapes and natural frequencies of more complex rooms can be determined from elaborate models such as the finite element method [Aba17], however in the sound zones problem they are usually determined from measurements [Møl+19].

Towards higher frequencies the modal density, i.e. the number of eigenfrequencies per bandwidth, increases and more modes are active at the same time. This results in a more complex sound field where it can be advantageous to rely on a statistical rather than the described analytical model [JJ13]. However, since the focus of this thesis is on low frequencies in medium sized rooms only the analytical model is used to simulate the sound field.

Using Green's function one can express the room transfer function $h_{ml,r}$ from equation 2.13 as

$$h_{ml,r} = j\omega\rho_0 G_{ml}. \tag{2.17}$$

and the resulting sound pressure as

$$p_m = j\omega\rho_0 q_l G_{ml}. \tag{2.18}$$

At 23° C in normal room conditions the speed of sound is $c = 345\ \frac{\text{m}}{\text{s}}$. The frequency range of interest in this report ranges from 20 Hz to 300 Hz, so the corresponding wavelengths range from more than 17 m at 20 Hz to around 1.15 m at 300 Hz. A source that is small compared to the wavelength can be approximated by a point source [JJ13], so it is fair to simulate loudspeakers as point sources in this frequency range. So the sound pressure at every location $\boldsymbol{r}_m$ in a room can be simulated by equation 2.18 when a loudspeaker emits the volume velocity q_l at a location $\boldsymbol{r}_l$.

An example of a sound field excited by a point source in a room is illustrated in figure 2.2. The speaker is positioned in the bottom left corner and the resulting sound field is simulated at every position in the room using Green's function. The sound field is analysed on the floor of the room, so $z = 0$. At 30 Hz the wavelength of the sound is nearly twice as long as the room dimension in the vertical y direction. Therefore the sound field resembles a standing wave pattern with very low sound pressure half way along the y axis and only very little changes across the x axis. At higher frequencies one can notice that the shorter wavelengths and the greater modal overlap cause the sound field to be more complex.

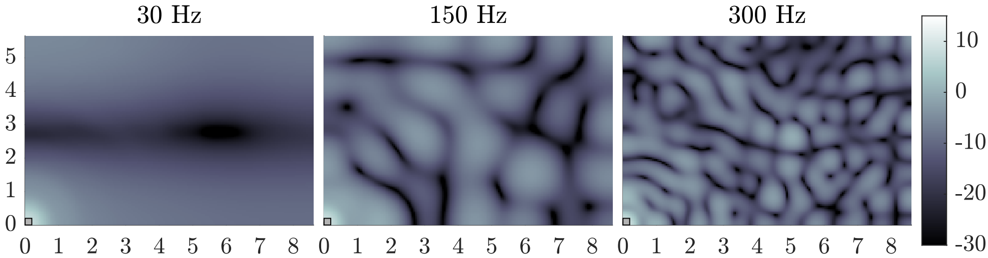


Figure 2.2: Magnitude of sound field in dB SPL when room is excited by a loudspeaker in the bottom left corner. The volume velocity of the source was set proportional to $\frac{1}{\omega}$ such that the SPL at different frequencies is proportional to Green's function. The ticks on the x and y axis denote the room dimensions in meters

2.2 Interference and the sound zones problem

2.2.1 Interference

The previous section introduced how a single source can excite a sound field in a room. A single source however has only two tunable parameters: the magnitude and the phase of the signal controlling the volume velocity q_l. Since the resulting sound pressure in equation 2.13 is a product of the volume velocity q_l and the response of the space $h_{ml,r}$ it becomes evident that one can not alter the spatial properties of a sound field, but can only change the phase and magnitude uniformly across the room. To create sound zones a certain degree of control is required that can be achieved from leveraging interference from a higher number of loudspeakers.

Interference is a fundamental phenomenon in acoustics that occurs when two or more sound waves encounter. Due to the principal of superposition interfering waves will add up linearly [Ros07]. Since the sound pressure is a complex quantity multiple sound waves can interfere destructively or constructively depending on the phase of the individual summands. If two waves are out of phase by more than $\pm 90°$ they will interfere destructively, when the phase angle is less than $\pm 90°$ they will interfere constructively. Constructive intereference implies that the summed sound pressure amplitude is greater than that of the greatest summand, destructive interference results in a sound pressure smaller than the greatest summand. Waves of equal amplitude can obtain complete destructive interference when 180° out of phase, i.e. the sound pressure will be canceled out completely, or complete constructive interference when perfectly in phase, i.e. the resulting amplitude is doubled [Ros07].

2.2.2 The sound zones problem

Now imagine the sound field generated by L loudspeakers being sampled by M microphones in a predefined dark and bright zone ($\frac{M}{2}$ per zone). The sound pressure at each microphone m is then the sum of the individual contributions from each speaker l, i.e.

$$p_m = \sum_{l=1}^{L} h_{ml,r} q_l = \sum_{l=1}^{L} h_{ml,r} h_{l,s} w_l x. \tag{2.19}$$

where w_l is the filter applied to the input signal x. Note that all variables p_m, h_{ml}, w_l, $h_{l,s}$ x and q_l are frequency dependent, but for the sake of readability the frequency dependent notation was omitted. In the time domain the product in equation 2.19 would be a convolution.

The signal x can be seen as an independent scaling factor. A sound zones system is designed to work for any input signal. So the signal x is neglected in the following. Equation 2.19 can then be expressed for all M microphones in matrix notation to

$$\boldsymbol{p} = \boldsymbol{H}_r(\boldsymbol{h}_s \odot \boldsymbol{w}) \tag{2.20}$$

where $\odot$ denotes element-wise multiplication and

$$\boldsymbol{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_M \end{bmatrix}, \; \boldsymbol{H}_r = \begin{bmatrix} h_{11,r} & h_{12,r} & \cdots & h_{1L,r} \\ h_{21,r} & h_{22,r} & \cdots & h_{2L,r} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M1,r} & h_{M2,r} & \cdots & h_{ML,r} \end{bmatrix}, \; \boldsymbol{h}_s = \begin{bmatrix} h_{1,s} \\ h_{2,s} \\ \vdots \\ h_{L,s} \end{bmatrix}, \; \boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix}.$$

By multiplying the speaker transfer function $\boldsymbol{h}_s$ with the columns of the room transfer function matrix $\boldsymbol{H}_r$ one can simplify equation 2.20 to

$$\boldsymbol{p} = \boldsymbol{H}\boldsymbol{w}, \tag{2.21}$$

where $\boldsymbol{H}$ is the transfer function matrix given by

$$\boldsymbol{H} = \boldsymbol{H}_r \odot (\boldsymbol{I}\boldsymbol{h}_s) = \begin{bmatrix} h_{11,r}h_{1,s} & h_{12,r}h_{2,s} & \cdots & h_{1L,r}h_{L,s} \\ h_{21,r}h_{1,s} & h_{22,r}h_{2,s} & \cdots & h_{2L,r}h_{L,s} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M1,r}h_{1,s} & h_{M2,r}h_{2,s} & \cdots & h_{ML,r}h_{L,s} \end{bmatrix}. \tag{2.22}$$

Here, $\boldsymbol{I}$ is the identity matrix of shape $M \times 1$. The transfer function matrix $\boldsymbol{H}$ consists of the full transfer functions from the input voltage at each of the L loudspeakers to the sound pressure at each of the M microphones.

Goal of the sound zones problem is to design a set of filters $\boldsymbol{w}$ that modify the phase and magnitude of the volume velocities of multiple sound sources such that the resulting pressure field will interfere destructively in a predefined dark zone and constructively in a bright zone.

To attain the filters w_l for each loudspeaker certain cost functions such as the *acoustic contrast* (AC) or the *reproduction error* can be optimized. The corresponding optimization methods are *acoustic contrast control* (ACC) [CK02] and *pressure matching* (PM) [KN93] respectively.

2.2.3 Acoustic Contrast Control

In ACC the acoustic contrast, i.e. the power ratio between the bright and the dark zone, is maximized. Defining the bright zone as the sound pressure p_{m_B} sampled by M_B microphones and the dark zone as the sound pressure p_{m_D} at M_D microphones one can formulate the acoustic contrast between the zones as [Møl+19]

$$contrast = \frac{{M_B}^{-1} \sum_{m_B}^{M_B} ||p_{m_B}||_2^2}{{M_D}^{-1} \sum_{m_D}^{M_D} ||p_{m_D}||_2^2}, \tag{2.23}$$

where $||\cdot||_2$ denotes the l_2-norm. The optimal acoustic contrast filters maximize the contrast, so they can be obtained from maximizing the cost function [Møl+19]

$$J_{acc}(\boldsymbol{w}) = \frac{{M_B}^{-1}||\boldsymbol{H}_B\boldsymbol{w}||_2^2}{{M_D}^{-1}||\boldsymbol{H}_D\boldsymbol{w}||_2^2}. \tag{2.24}$$

This can be seen as an eigenvalue problem with the solution [Møl+19]

$$\boldsymbol{w}_{acc} = \theta\Big(\big({\boldsymbol{H}_D}^H \boldsymbol{H}_{\boldsymbol{D}}\big)^{-1} {\boldsymbol{H}_B}^H \boldsymbol{H}_B\Big), \tag{2.25}$$

where the operator $\theta(\cdot)$ denotes a function returning the eigenvector corresponding to the largest eigenvalue of its input and $(\cdot)^H$ denotes the Hermitian transpose of a matrix.

2.2.4 Pressure Matching

In pressure matching the filters are designed such that the resulting sound pressure minimizes the *reproduction error*, i.e. the mean squared error to a predefined target sound field $\boldsymbol{p}_t$ [KN93]. The target sound field is designed such that the pressure in the bright zone resembles a desired sound field (e.g. a plane wave) and zero pressure in the dark zone. The optimal pressure matching filters can be found by minimizing the cost function [KN93]

$$J_{pm}(\boldsymbol{w}) = ||\boldsymbol{p}_t - \boldsymbol{H}\boldsymbol{w}||_2^2. \tag{2.26}$$

This minimization is a Least Squares problem, which is discussed in further detail in section 2.4 and to which a solution can be found by [GL13; The15; KN93]

$$\boldsymbol{w}_{PM} = (\boldsymbol{H}^H\boldsymbol{H})^{-1}\boldsymbol{H}^H\boldsymbol{p}_t. \tag{2.27}$$

An example of a sound field generated by PM is illustrated in figure 2.3. One can notice how the pressure matching filters manipulated phase and magnitude such that the resulting superimposed sound field from the eight loudspeakers is very low in the dark zone relative to the SPL in the bright zone.

Though PM generally does not achieve the same contrast as ACC [Col+14a] it provides more control about the sound field in the bright zone. In ACC the wavefronts in the bright zone come from erratic directions [Jac+11], which was found to reduce sound quality in perceptual studies [Bay+15] compared to plane waves. This problem can be overcome by using Planarity Control [Col+14a], which imposes an additional constraint on the ACC solution such that the sound field in the bright zone resembles a plane wave.

To improve the acoustic contrast from PM the transfer function matrix $\boldsymbol{H}$ can be weighted such that the reduction of sound energy in the dark zone gains significance compared to the mean squared error in the bright zone [CJ13]. The PM solution will approach the ACC solution when the dark zone transfer functions are weighted heavily [CJ13], so ACC can be seen as a weighted version of PM. Therefore this thesis focuses on PM in the following, which allows for more control on the sound field.

2.3 Challenges in the sound zones problem

In situ the sound zones problem is usually designed as a feed forward system. While simulations such as in figure 2.3 are based on the exact underlying transfer functions these are not available when designing a sound zones system in e.g. a domestic room. The transfer functions have to be determined from measurements that will contain noise, therefore the transfer function matrix used to calculate the filters is only an estimate of the real underlying transfer functions [Møl+19].

Another problem arises from the fact that transfer functions depend on underlying physical properties of the room that can change over time. The room itself can change, e.g. when a window is opened, persons move around in the room or furniture is changed. However this thesis assumes the room to be static and since the wavelengths in the observed frequency range are long compared to e.g. a human head the effect of scattering is assumed to be small [CJ12]. This thesis focuses on

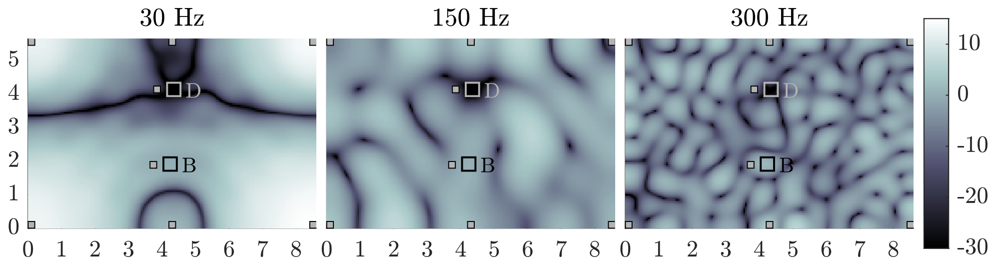


Figure 2.3: Sound field generated by eight loudspeakers standing on the floor of a 5.65 m× 8.61 m× 2.7 m room, when the filters are calculated using PM. The loudspeaker positions are indicated by small squares. The bright and dark zone are both 50 cm× 50 cm× 30 cm and are indicated with B and D respectively. The sound field is visualized at a height of 4.14 m, in the middle of the sound zones. The color bar indicates the SPL in dB.

changes in the room transfer functions $h_{ml,r}$ due to temperature changes and changes in the speaker transfer functions $h_{l,s}$. In an earlier work [Ben20] it was shown that changes in temperature and speaker transfer function can have a significant impact on the performance of the system, since the sound zones solution is based on a specific transfer function matrix that does not represent the real transfer function matrix.

2.3.1 Changes in Speaker Transfer Functions

Both, non-linear distortion products as well as changes in the transfer functions can decrease the performance of sound zone systems. Ma et al. [Ma+18; Ma+19] investigated the impact of loudspeaker distortion on the sound zones problem. It was shown that effects of non-linear distortion can sometimes be audible in the dark zone, but the effect is small compared to the effects caused by deviations between measured and real transfer functions [Ma+18]. The impact of the distortion can further be reduced by using high quality loudspeakers and reducing the control effort [Ma+19], i.e. the signal power relative to a reference [CK02]. This thesis therefore focuses on the effect of changes in the linear frequency response of the speakers.

One cause of changes in the frequency response results from the voice coil getting hot because the loudspeaker is used for a long time at high levels. The DC-Resistance R_E of a wire made from copper changes with temperature as [Kas05]

$$\tilde{R}_E = R_E(1 + \alpha(T - T_0)), \tag{2.28}$$

where $\alpha = 4.04 \cdot 10^{-3}\ \mathrm{K}^{-1}$ is the temperature coefficient of copper, $T_0 = 23\ °\mathrm{C}$ is the reference temperature and $\tilde{R}_E$ is the DC resistance at the temperature T. Another effect of the coil heating up arises when the heat dissipates to the suspension of the coil, i.e. the spider, which gets softer with rising temperatures [PA07]. The compliance of the diaphragm further increases (gets softer) as it heats up due to mechanical losses [PA07] and due to the mechanical stretching and squeezing of the oscillation [SK01]. The effect of the temperature on the force factor Bl is assumed to be small [Age07]. The moving mass M_{AT} and the box compliance C_{AB} and losses R_{AB} are not changing. While it is possible that the losses in the suspension R_{AS} can change, the effect is assumed to be small since these losses are much smaller than the electrical losses R_{AE} and the losses in the box R_{AB}.

Consequently this report focuses on the changes in the DC-Resistance of the coil R_E and the change in suspension compliance C_{AS}. From equations 2.8 - 2.10 it can be seen that an increase in DC-Resistance will increase damping, the Q-Factor Q_{TC} will decrease, and the overall emitted volume velocity will be smaller. A higher suspension compliance C_{AS} will decrease the resonance frequency and the Q-Factor.

2.3.2 Changes in room temperature

The speed of sound is proportional to the square root of the absolute temperature of the medium [JJ13]. At 20 °C the speed of sound can be calculated to $c_{20^\circ} = 343\ \frac{m}{s}$

and to $c_{26^\circ} = 347\ \frac{m}{s}$ at 26 °C. The wavenumber k is inversely proportional to the speed of sound. Since the time constant τ_ν is inversely proportional to c^2 [JJ13] in this simplified model, the damping term $-jk/(\tau_\nu c)$ is independent of the speed of sound. A change in speed of sound will therefore result in Green's function from equation 2.15 to be stretched or squeezed along the frequency axis for a decrease or increase in temperature respectively.

This effect is further illustrated in the example of a rectangular room in section 3.3, where one of the transfer functions used for the following simulations is shown.

2.4 The Least Squares problem

Pressure Matching is a Least Squares problem. The following section defines the Least Squares problem and shows how the solution can be derived and further introduces a measure of sensitivity of the solution.

2.4.1 Formulating the Least Squares problem

Given a data matrix $\boldsymbol{A} \in \mathbb{C}^{M\times L}$ and a vector of observations $\boldsymbol{b} \in \mathbb{C}^M$, the least squares problem aims at solving the linear system

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{2.29}$$

with the smallest possible error by finding the vector $\boldsymbol{x}_{LS} \in \mathbb{R}^L$ that minimizes $||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||^2$ with respect to $\boldsymbol{x}$, where $||\cdot||_2$ denotes the l_2-norm. So the LS problem can be expressed as

$$\boldsymbol{x}_{LS} = \underset{\boldsymbol{x}}{\operatorname{argmin}}\Big\{||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||_2\Big\}. \tag{2.30}$$

If $M > L$ then the system is overdetermined, i.e. there are more equations than unknowns. If $\boldsymbol{A}$ is of full rank, then there exists a unique $\boldsymbol{x}_{LS}$ minimizing equation 2.30. A matrix is said to be of full rank when its rank r is equal to the largest possible for a matrix of the same dimensions, which is the lesser of the number of rows and columns M and L. For overdetermined systems the rank of $\boldsymbol{A}$ is full when $r = L$.

If $\boldsymbol{b}$ is in the span of $\boldsymbol{A}$, then $\boldsymbol{b}$ can be expressed by a linear combination of the column vectors of $\boldsymbol{A}$, so there exists a vector $\boldsymbol{x}$, such that $||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||_2 = 0$. However, for overdetermined systems $\boldsymbol{b}$ is usually not in the span of $\boldsymbol{A}$, so $||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||_2 > 0, \forall \boldsymbol{x}$.

The product $\boldsymbol{A}\boldsymbol{x}_{LS}$ that comes closest to $\boldsymbol{b}$ in the l_2-norm is the projection of $\boldsymbol{b}$ on the column space of $\boldsymbol{A}$, so the difference vector $\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_{LS}$ will be orthogonal to the column space of $\boldsymbol{A}$. The column space of $\boldsymbol{A}$ equals the row space of $\boldsymbol{A}^H$, so multiplying $\boldsymbol{A}^H$ with the difference vector $\boldsymbol{r}$ yields

$$\boldsymbol{A}^H(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_{LS}) = 0 \tag{2.31}$$

due to the orthogonality. Reformulating leads to the ordinary least squares solution (OLS) which can be expressed as

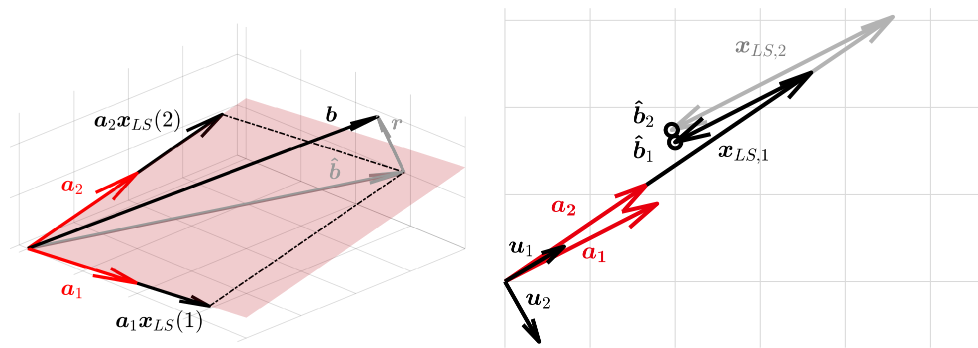
$$\boldsymbol{x}_{LS} = \left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H\boldsymbol{b}. \tag{2.32}$$

This equation (2.32) allows us to find the optimal linear combination $\boldsymbol{x}_{LS}$ to come as close as possible to $\boldsymbol{b}$ in the span of $\boldsymbol{A}$. The projection $\hat{\boldsymbol{b}}$ of $\boldsymbol{b}$ on the span of $\boldsymbol{A}$ is illustrated in figure 2.4a.

If the rank of $\boldsymbol{A}$ is less than L, then there is an infinite number of solutions to the LS problem, as the columns of $\boldsymbol{A}$ are not linearly independent, such that there are infinitely many ways to represent the projection of $\boldsymbol{b}$ in the span of $\boldsymbol{A}$. This makes the calculation of equation 2.32 impossible, as the inverse of a rank deficient matrix does not exist and $rank(\boldsymbol{A}^H\boldsymbol{A}) \leq rank(\boldsymbol{A})$.

The ordinary least squares solution can also be expressed using the singular value decomposition $\boldsymbol{A} = \boldsymbol{U\Sigma V}$. The left singular vectors $\boldsymbol{u}_i$, i.e. the columns of $\boldsymbol{U}$, form an orthonormal basis spanning the same space as the span of $\boldsymbol{A}$. So projecting $\boldsymbol{b}$ on the span of $\boldsymbol{A}$ will result in the same projection $\hat{\boldsymbol{b}}$ as when projecting $\boldsymbol{b}$ on the vectors $\boldsymbol{u}_i$. This projection can be expressed as

$$\hat{\boldsymbol{b}} = \sum_{i=1}^{r}(\boldsymbol{u}_i{}^T\boldsymbol{b})\boldsymbol{u}_i. \tag{2.33}$$



(a) Projection $\hat{\boldsymbol{b}}$ of vector $\boldsymbol{b}$ on the surface spanned by the column vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ of the matrix $\boldsymbol{A}$. The residual $\boldsymbol{r}$ is orthogonal to every vector in the spanned space. The projection $\hat{\boldsymbol{b}}$ equals the product of $\boldsymbol{A} = [\boldsymbol{a}_1,\ \boldsymbol{a}_2]$ and $\boldsymbol{x}_{LS} = [\boldsymbol{x}_{LS}(1),\ \boldsymbol{x}_{LS}(2)]^T$

(b) Sensitivity of the OLS solution when the basis vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are collinear. Small perturbations in the projection $\hat{\boldsymbol{b}}_1$ can cause drastic changes in the solution $\boldsymbol{x}_{LS,1}$. $||\boldsymbol{x}_{LS,2}||_2$ is significantly greater than $||\boldsymbol{x}_{LS,1}||_2$ even though the projection $\hat{\boldsymbol{b}}_2$ is close to $\hat{\boldsymbol{b}}_1$. $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are the left singular vectors of $\boldsymbol{A}$

Figure 2.4: Sketches to illustrate LS

This projection can then be used to find an alternative but equal formulation to equation 2.34. With $\boldsymbol{v}_i$ representing the right singular vectors, i.e. the columns of $\boldsymbol{V}$, and the singular values σ_i the OLS can be formulated to

$$\boldsymbol{x}_{LS} = \sum_{i=1}^{r} \frac{\boldsymbol{u}_i{}^T \boldsymbol{b} \boldsymbol{v}_i}{\sigma_i}. \tag{2.34}$$

Formulating the least squares solution in terms of the singular value decomposition can be advantageous when analysing the system on the sensitivity to perturbations.

2.4.2 Sensitivity of the LS solution

The least squares solution from equation 2.32 is very sensitive to small perturbations if either $\boldsymbol{b}$ is nearly orthogonal to the span of $\boldsymbol{A}$ or if the system is *ill-conditioned*.

If $\boldsymbol{b}$ is nearly orthogonal to the span of $\boldsymbol{A}$, then the projection of $\boldsymbol{b}$ onto the span of $\boldsymbol{A}$ will be much smaller than the vector $\boldsymbol{b}$ itself. It follows that relatively small changes in $\boldsymbol{b}$ can result in big changes in the LS solution $\boldsymbol{x}_{LS}$. However, $\boldsymbol{b}$ being almost orthogonal to the span of $\boldsymbol{A}$ also implies a very poor LS solution in general, making this scenario less relevant for further investigation.

A linear system is said to be *ill-conditioned* when small changes in either the matrix $\boldsymbol{A}$ or in the vector $\boldsymbol{b}$ result in great changes in the least squares solution. From a geometrical point of view this is the case if the vectors in the span of $\boldsymbol{A}$ are nearly linear dependent as illustrated in figure 2.4b. A metric that evaluates the sensitivity of the least-squares problem is the condition number. In the l_2-norm the condition number of a matrix $\boldsymbol{A}$ is defined as [GL13]

$$\kappa(\boldsymbol{A}) = \frac{\sigma_{max}}{\sigma_{min}}, \tag{2.35}$$

where σ_{max} and σ_{min} are the maximum and minimum singular values of $\boldsymbol{A}$. When solving the LS problem using equation 2.32, the condition number of $\boldsymbol{A}^H\boldsymbol{A}$ is critical and can be calculated by [GL13]

$$\kappa(\boldsymbol{A}^H\boldsymbol{A}) = ||\boldsymbol{A}^H\boldsymbol{A}||_2||(\boldsymbol{A}^H\boldsymbol{A})^{-1}||_2 = \kappa(\boldsymbol{A})^2. \tag{2.36}$$

A matrix with linear dependent columns is not invertible, the lowest singular value will be zero, so the condition number is infinite. The condition number for matrices with highly correlated columns will be high. An orthonormal matrix will have a condition number of 1, since all singular values are 1. The condition number can therefore also be seen as a metric evaluating the orthonormality of the matrix columns.

From an analysis of the SVD decomposition of the OLS, equation 2.34, it can be observed that the solution $\boldsymbol{x}_{LS}$ is most sensitive when perturbations in $\boldsymbol{b}$ occur in the directions $\boldsymbol{u}_i$ corresponding to low singular values σ_i. The OLS will therefore be very robust if the data matrix is orthonormal and all singular values are one and can become very sensitive if the low singular values are small. This also explains the high sensitivity to the small perturbation in figure 2.4b. Both vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are pointing in a similar direction. Therefore the singular value σ_1 corresponding to the first singular vector $\boldsymbol{u}_1$ will be much bigger than the second singular value σ_2. The perturbation $\hat{\boldsymbol{b}}_1 - \hat{\boldsymbol{b}}_2$ happens in the direction of the second singular vector. If the perturbation had been along the first singular value the solution would only have changed little.

The projections $\hat{\boldsymbol{b}}_1$ and $\hat{\boldsymbol{b}}_2$ in figure 2.4b are close to the axis $\boldsymbol{u}_1$. Whenever the projections are further away from this axis the l_2-norm of $\boldsymbol{x}_{LS}$ can easily become very large. For large solutions $\boldsymbol{x}_{LS}$ the system becomes sensitive to small changes in $\boldsymbol{A}$, explaining why collinearity is not just causing sensitivity to perturbations in $\hat{\boldsymbol{b}}$ but also in $\boldsymbol{A}$.

2.5 Regularization

All measured data contains noise and whenever there is noise in the system it is usually preferred to find a robust, rather than the optimal solution for the measured system. To find a more robust solution one can make use of regularization.

2.5.1 Ridge regression

A common method of finding a more general LS solution to linear systems is ridge regression. Ridge regression applies a special kind of Tikhonov regularization, where the regularization matrix is chosen to be a multiple of the identity matrix $\boldsymbol{I}$. Ridge regression can be formulated by the minimization problem

$$\boldsymbol{x}_{RR} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \Big\{ ||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||_2^2 + \lambda ||\boldsymbol{x}||_2^2 \Big\}, \tag{2.37}$$

where λ is a design parameter that controls the scale of the regularization with the constraint $\lambda \geq 0$. When $\lambda = 0$ the problem reduces to the ordinary least squares problem without regularization from equation 2.30. Great values of λ will enforce great regularization to the problem, the solution will fit the given system less accurately but will also avoid overfitting. The minimization problem is solved by

$$\boldsymbol{x}_{RR} = (\boldsymbol{A}^H \boldsymbol{A} + \lambda \boldsymbol{I})^{-1} \boldsymbol{A}^H \boldsymbol{b}. \tag{2.38}$$

Ridge regression applies a weight to the l_2-norm of $\boldsymbol{x}$. Thereby it prevents the solution from growing too big, making the solution less sensitive to noise. It can also be understood as a diagonal weight to the Gram matrix $\boldsymbol{A}^H\boldsymbol{A}$. The Gram matrix is a symmetric matrix containing the dot products of the different column vectors of the matrix $\boldsymbol{A}$. For an orthogonal $\boldsymbol{A}$ the product $\boldsymbol{A}^H\boldsymbol{A}$ will be diagonal. The closer the span of $\boldsymbol{A}$ is to linear dependence, the greater the corresponding non-diagonal elements in $\boldsymbol{A}^H\boldsymbol{A}$ will be. Adding a diagonal loading $\lambda\boldsymbol{I}$ to the matrix product will diagonalize the matrix, making it closer to orthogonality and thereby reduces the condition number and increases numeric stability.

A similar, interesting interpretation of ridge regression can be obtained by leveraging the singular value decomposition. By applying the SVD to the ridge regression solution 2.38 it can be shown [The15] that the regularized projection of $\boldsymbol{b}$ onto the span of $\boldsymbol{X}$, i.e. $\hat{\boldsymbol{b}}$, can be expressed as

$$\hat{\boldsymbol{b}} = \sum_{i=1}^{r} \frac{{\sigma_i}^2}{\lambda + {\sigma_i}^2}(\boldsymbol{u}_i^T \boldsymbol{b})\boldsymbol{u}_i. \tag{2.39}$$

So the ridge regression solution is

$$\boldsymbol{x}_{RR} = \sum_{i=1}^{r} \frac{\sigma_i}{\lambda + {\sigma_i}^2}({\boldsymbol{u}_i}^T \boldsymbol{b})\boldsymbol{v}_i. \tag{2.40}$$

It is interesting to note that ridge regression assigns higher weights to more informative directions where most of the data variance takes place and regularizes directions corresponding to low singular values. Since the ordinary least squares solution is sensitive to noise in the directions corresponding to low singular values, ridge regression can be very effective if the sensitivity of the problem is explained by linear dependence of the column vectors. If the column vectors have low collinearity the singular values will all be close to one, making the system less sensitive to perturbations and ridge regression less effective.

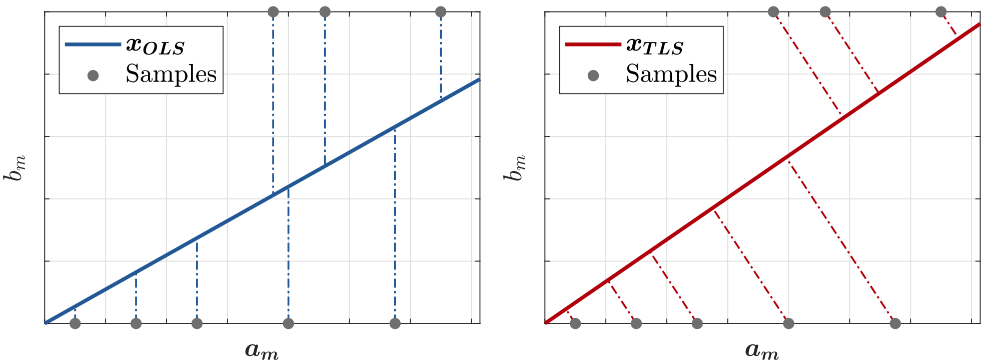
2.5.2 Total Least Squares

The ordinary least squares solution from equation 2.32 finds the hyperplane that minimizes the sum of squared residuals in the observations. This implies that the regressor $\boldsymbol{A}$ is exact. However, since the regressor $\boldsymbol{A}$ is often based on noisy measurements it can be advantageous to find a solution that allows for errors in both, the regressor and the observations. The total least squares solution finds the hyperplane that minimizes the total square distance to all samples $[\boldsymbol{a}_m,\ b_m]$ from it by

$$\begin{aligned}
\boldsymbol{x}_{TLS} &= \underset{\boldsymbol{x}}{\operatorname{argmin}} \Big\{ ||\hat{\boldsymbol{A}} - \boldsymbol{A},\ \hat{\boldsymbol{b}} - \boldsymbol{b}||_F^2 \Big\} \\
&= \underset{\boldsymbol{x}}{\operatorname{argmin}} \Big\{ \sum_{m=1}^{M} (\hat{b}_m - b_m)^2 + ||\hat{\boldsymbol{a}}_m - \boldsymbol{a}_m||_F^2 \Big\} \\
&= \underset{\boldsymbol{x}}{\operatorname{argmin}} \Big\{ \sum_{m=1}^{M} (\hat{\boldsymbol{a}}_m \boldsymbol{x} - b_m)^2 + ||\hat{\boldsymbol{a}}_m - \boldsymbol{a}_m||_F^2 \Big\},
\end{aligned} \tag{2.41}$$

such that all $[\hat{\boldsymbol{a}}_{\boldsymbol{m}},\ \hat{b}_m]$ lie on that hyperplane. The difference between OLS and TLS is illustrated in figure 2.5.

To arrive at the TLS solution one can consider the column space of $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{b}}$. Since $\hat{\boldsymbol{b}}$ is a linear combination of the columns of $\hat{\boldsymbol{A}}$ it can be inferred that the rank of the concatenated matrix $[\hat{\boldsymbol{A}},\ \hat{\boldsymbol{b}}]$ equals the rank r of the matrix $\hat{\boldsymbol{A}}$. If $\hat{\boldsymbol{A}}$ is a tall



(a) Ordinary least squares finds the hyperplane minimizing the sum of squares of the errors in the observations $\sum_{m=1}^{M}(\hat{b}_m - b_m)^2$

(b) Total least squares minimizes the total square distance of all samples to the hyperplane by minimizing $\sum_{m=1}^{M}(\hat{b}_m - b_m)^2 + ||\hat{\boldsymbol{a}}_m - \boldsymbol{a}_m||_F^2$

Figure 2.5: Ordinary least squares and total least squares will find different solutions given the same input samples $[\boldsymbol{a}_m,\ b_m]$. The dashed lines represent the respective error of which the sum of squares is minimized.

matrix of full rank then the rank will equal the number of columns L, so $r = L$. The concatenated matrix $[\boldsymbol{A},\ \boldsymbol{b}]$ is of rank $r = L + 1$ since $\boldsymbol{b}$ is usually not in the column space of $\boldsymbol{A}$. Now, to solve equation 2.41 one essentially wants to find the best rank L approximation $[\hat{\boldsymbol{A}},\ \hat{\boldsymbol{b}}]$ to the rank $L + 1$ matrix $[\boldsymbol{A},\ \boldsymbol{b}]$. Following Eckart-Young-Mirsky's theorem the best low rank approximation in the Frobenius norm can be achieved by a truncated singular value decomposition [The15]. So the best rank L approximation of $[\boldsymbol{A},\ \boldsymbol{b}]$ in the Frobenius norm is given by

$$[\hat{\boldsymbol{A}},\ \hat{\boldsymbol{b}}] = \sum_{i=1}^{L} \bar{\sigma}_i \bar{\boldsymbol{u}}_i \bar{\boldsymbol{v}}_i{}^T. \tag{2.42}$$

Where $[\boldsymbol{A},\ \boldsymbol{b}] = \sum_i^{L+1} \bar{\sigma}_i \bar{\boldsymbol{u}}_i \bar{\boldsymbol{v}}_i^T$ is the SVD of the concatenated sample matrix.

Using the SVD of the regressor $\boldsymbol{A} = \sum_i^L \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i$ one can compute the approximations $\hat{\boldsymbol{b}}$ by

$$\hat{\boldsymbol{b}} = \sum_{i=1}^{L} \frac{\sigma_i{}^2}{\sigma_i{}^2 - \bar{\sigma}_{L+1}^2} (\boldsymbol{u}_i{}^T \boldsymbol{b}) \boldsymbol{u}_i \tag{2.43}$$

and the TLS solution $\boldsymbol{x}_{TLS}$ to

$$\begin{aligned} \boldsymbol{x}_{TLS} &= \sum_{i=1}^{L} \frac{\sigma_i}{\sigma_i{}^2 - \bar{\sigma}_{L+1}^2} (\boldsymbol{u}_i^T \boldsymbol{b}) \boldsymbol{v}_i \\ &= (\boldsymbol{A}^H \boldsymbol{A} - \bar{\sigma}_{L+1}^2 \boldsymbol{I})^{-1} \boldsymbol{A}^H \boldsymbol{b}. \end{aligned} \tag{2.44}$$

The fact that the approximated concatenated sample matrix can be expressed by the SVD of the regressor has the implications that the hyperplane minimizing the total distance from all points is defined by $\boldsymbol{v}_{L+1}$ and that the squared error that is minimized in equation 2.41 is of the size of the last singular value $||\hat{\boldsymbol{A}} - \boldsymbol{A},\ \hat{\boldsymbol{b}} - \boldsymbol{b}||_F^2 = \bar{\sigma}_{L+1}$ [The15]. It is therefore possible to approximate $[\boldsymbol{A},\ \boldsymbol{b}]$ with little error, however this does not imply that the error in the predictions will be small since TLS leverages an optimal approximation of $\boldsymbol{A}$ to minimize the error $||\hat{\boldsymbol{A}} - \boldsymbol{A},\ \hat{\boldsymbol{b}} - \boldsymbol{b}||_F^2$. This optimal approximation will usually deviate from the real underlying sample matrix.

Comparing the TLS solution 2.44 to the ridge regression solution 2.38 one can notice the deregularizing property of the TLS. The TLS is in fact promoting a larger unstable solution [The15]. TLS will impose a higher condition number to the matrix that is to be inverted than the OLS or RR solution. So when both the regressor and the observations contain noise TLS can be superior to OLS and RR, but only when the noise is small [The15].

2.5.2.1 Regularized Total Least Squares

To compensate the high noise sensitivity of the TLS regularization can be added to the problem. The regularized solution can then be written as [HO96]

$$\boldsymbol{x}_{RTLS} = (\boldsymbol{A}^H\boldsymbol{A} - \bar{\sigma}_{L+1}^2\boldsymbol{I} + \lambda\boldsymbol{I})^{-1}\boldsymbol{A}^H\boldsymbol{b}. \tag{2.45}$$

Comparing this equation to the OLS solution 2.32 and the RR solution 2.38 it can be observed that setting $\lambda = \bar{\sigma}_{L+1}^2$ will result in the OLS and setting $\lambda > \bar{\sigma}_{L+1}^2$ will result in the RR solution. If $\lambda < 0$ the RTLS solution will be even more sensitive to errors [HO96]. So the most interesting range of λ is between 0 and $\bar{\sigma}_{L+1}^2$.

2.5.3 Specific adaptation in the sound zones problem

Both TLS and RR assume the presence of additive noise in the observations or the regressor. The goal of these regularization techniques is to make the system robust to this additive noise. However, in many situations the perturbations in the system occur due to a known change in the physical environment and it can be advantageous to leverage this knowledge to adapt the system, rather than simply making the system robust to additive noise. In practice the underlying physical environment and the corresponding changes are often not perfectly understood or it is not possible to monitor the changes exactly, so the gain of specific adaptation is limited.

2.5.3.1 Resampling Temperature

In the sound zones problem the observation matrix is the matrix of transfer functions. The scenario in which we assume the deviations in the transfer function matrix to be caused by a change in temperature and the resulting speed of sound difference is special, as the real transfer functions can be determined when assuming the temperature in the room to be uniform. It was argued in section 2.3.2 that a change in temperature causes the transfer functions to stretch or squeeze along the frequency axis when the temperature is decreased or increased respectively. So assuming the boundary conditions in the room to be independent of temperature and the temperature change in the room to be uniform enables us to update our transfer function estimates from the initial estimates by resampling. Since the change in the transfer functions is proportional to the change in speed of sound, which in turn is proportional to the square root of the relative change between the new temperature T_{new} and the initial temperature T_{init}, the updated transfer functions are given by

$$h_{ml,new}\Big(f\Big) = h_{ml}\Big(\sqrt{\frac{T_{new}}{T_{init}}}f\Big) \tag{2.46}$$

and the updated filters can be expressed in the same fashion as

$$w_{l,new}\Big(f\Big) = w_{l}\Big(\sqrt{\frac{T_{new}}{T_{init}}}f\Big). \tag{2.47}$$

2.5.3.2 Probability model optimization (PMO)

Whenever the errors in the transfer function matrix underlie a known distribution one can leverage this knowledge and incorporate it in the cost functions [Møl+19]. The idea of PMO in pressure matching is to minimize the expected reproduction error, rather than minimizing the reproduction error for an initial transfer function matrix $\boldsymbol{H}_0$ [Zhu+17]. Zhu et al. [Zhu+17] formulate PMO filters for additive and multiplicative errors with normal or uniform distributions. In this thesis however, the PMO filters are designed to minimize the expected reproduction error for any error model independent of the error type. The filters that minimize the expected reproduction error can be found by

$$\boldsymbol{w}_{PMO} = \big(\mathbb{E}(\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}})\big)^{-1}\mathbb{E}(\tilde{\boldsymbol{H}})\boldsymbol{p}_t, \tag{2.48}$$

where $\tilde{\boldsymbol{H}}$ is the stochastic transfer function matrix with known means and error distributions and the operator $\mathbb{E}(\cdot)$ denotes the expected value.

CHAPTER 3

Simulation Method

This chapter introduces the procedure and the set of parameters in the simulations used to estimate the performance of a sound zones system. First the loudspeaker model is introduced that estimates reasonable transfer functions and deviations of their initial state.

3.1 General Setup

The sound zones problem is modeled in a perfectly rectangular room which is 5.65 m long, 8.61 m wide and 2.7 m high. The room has a reverberation time of 0.6 seconds in the whole frequency range from 20-300 Hz. Eight speakers are placed on the floor of the room. Four in the four corners, two half way along the width of the room by the walls and two half a meter behind the sound zones in the middle of the room as illustrated in figure 2.3. The dark and the bright zone are both $50 \times 50 \times 30$ cm and sampled by 75 microphones in each zone with 10 cm spacing between each. Therefore the transfer function matrix $\boldsymbol{H}$ has $M = 150$ rows and $L = 8$ columns. Each element h_{ml} is the product of the speaker transfer function $h_{l,s}$ and the room transfer function $h_{ml,r}$ as in equation 2.22.

To obtain the pressure matching solution a target sound field $\boldsymbol{p}_t$ has to be defined. Here the target sound field was set to equal the transfer functions from the closest speaker in the bright zone and zero in the dark zone. This ensures that the target sound field is not subject to phase issues which have shown to decrease the sound quality significantly [Bay+15]. The goal is therefore to create a sound field in the bright zone which is perceived as if only one loudspeaker was used and all the other loudspeakers are used to cancel the sound pressure in the dark zone. Using superposition this procedure could be performed for all eight speakers to enable a surround sound experience.

The following simulations incorporate random parameters, namely the room temperature and the voice-coil temperature. To model the influence of these parameters a Monte-Carlo simulation was performed. Each of the following simulations was performed with 200 Monte-Carlo iterations. A convergence study was performed for each simulation to verify that the simulation had converged well before the 200 iterations.

All calculations were executed in *MATLAB R2019b*. The corresponding code can be found on *GitLab*[1].

[1] `https://gitlab.gbar.dtu.dk/s190247/thesis`

3.2 Modelling the speaker transfer function

The model used to design the transfer functions of the speakers is based on the closed-box model by Leach et al. [LL03]. The model and the corresponding *MATLAB* implementation are based on equation 2.8 from section 2.1.1. A list of parameters is necessary to design a realistic model. As described in section 2.3.1 some of the speaker parameters can change when the speakers are used for a long time as the coil and the suspension heat up and thereby change the DC-Resistance of the voice coil R_E and the suspension compliance C_{AS} of the suspension. In the following the cold, initial state of the loudspeaker model is introduced before an estimation is made about expected changes in the transfer functions when the speaker heats up.

3.2.1 Initial parameters

The driver parameters, i.e. the DC-Resistance R_E of the coil, the force factor Bl, the piston area of the diaphragm S_D, the total moving mass of the coil diaphragm assembly M_{MD}, the suspension compliance of the diaphragm C_{MS} and the mechanical Q-Factor Q_{MS} were taken from the specification sheet of the *SLS-P830668* by *Tymphany HK Limited* [Lim17]. The mechanical impedances from the total moving mass of the coil diaphragm assembly M_{MD} and the suspension compliance C_{MS} were transformed to acoustic impedances by [LL03]

$$M_{AD} = \frac{M_{MD}}{{S_D}^2} \tag{3.1}$$

$$C_{AS} = C_{MS}{S_D}^2. \tag{3.2}$$

Further the acoustic resistance of the diaphragm modelling the losses in the suspension R_{AS} was not specified in the specifications of the driver and had to be derived from the mechanical Q-Factor [LL03]

$$R_{AS} = \frac{1}{Q_{MS}\sqrt{\frac{M_{AD}}{C_{AS}}}}. \tag{3.3}$$

The box was designed to have a volume V_B of 20 l. The effective acoustic volume V_{AB} of the box is usually slightly bigger to account for filling in the box [LL03]. Here the volume of the box is scaled by 1.15 to yield the effective acoustic volume V_{AB}. The effective acoustic volume of the box can then be used to determine the acoustic air-load compliance on the rear of the diaphragm to [LL03]

$$C_{AB} = \frac{V_{AB}}{\rho_0 c^2}. \tag{3.4}$$

Using a mass loading factor B and the effective density of the air and the filling in the box ρ, the acoustic moving mass M_{AB} inside the box can be determined by [LL03]

$$M_{AB} = \frac{B\rho}{\sqrt{S_D}\pi}. \tag{3.5}$$

The chosen values for ρ and B are based on examples provided by Leach [LL03].

The values used for the initial state transfer function are shown in table 3.1 and 3.2. Using these values the initial state total Q-Factor and closed-box resonance frequency $f_c = \omega_c/(2\pi)$ were determined to $Q_{TC,i} = 1.19$ and $f_{c,i} = 57$ Hz.

Effective piston area S_D	346.4 cm^2
Force factor Bl	10.42 Tm
Moving Mass M_{MD}	56.3 g
Mechanical Q-Factor Q_{MS}	7.15
Susp. compliance C_{MS}	429 µm/N
DC-Resistance Coil R_E	5.61 Ω

Table 3.1: Driver parameters [Lim17]

Box volume V_B	20 l
Eff. box volume V_{AB}	$1.15 \cdot 20$ l
Eff. density ρ	7.18 kg/m^3
Mass loading factor B	0.65

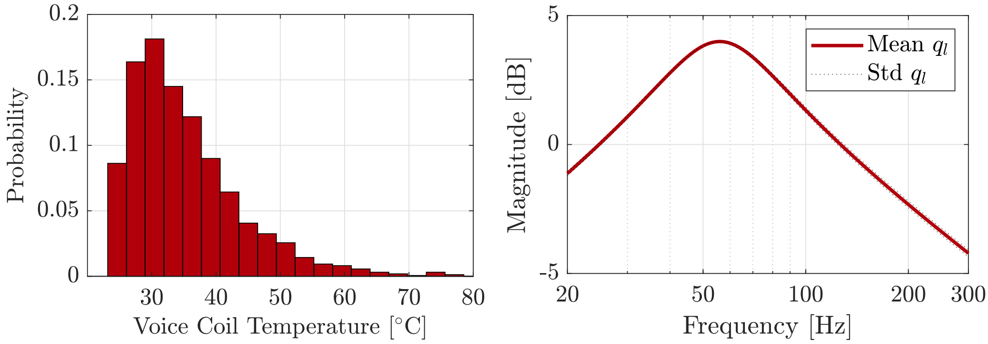
Table 3.2: Box parameters [LL03]

3.2.2 Deviations from the initial parameters

As explained in section 2.3.1 the greatest effect on the transfer function is expected from changes in the DC-Resistance R_E and the coil-diaphragm suspension C_{AS}. Both changes occur primarily when the speaker is driven at high levels, so it is reasonable to assume them to be correlated. Chapman introduced a model [Cha98] to simulate the voice coil temperature. This work includes measurements of the voice coil temperature while playing "The Dark Side of the Moon" by *Pink Floyd*, a 43 minute album with great loudness variation. A Gamma distribution that approximates the measured voice coil temperature distribution was employed in the Monte-Carlo simulation. The corresponding temperature histogram is shown in figure 3.1a. The resulting change in resistance was then modeled according to equation 2.28. Each speaker was modeled independently, i.e. no correlation between the voice coil temperatures of different speakers were assumed in this simulation.

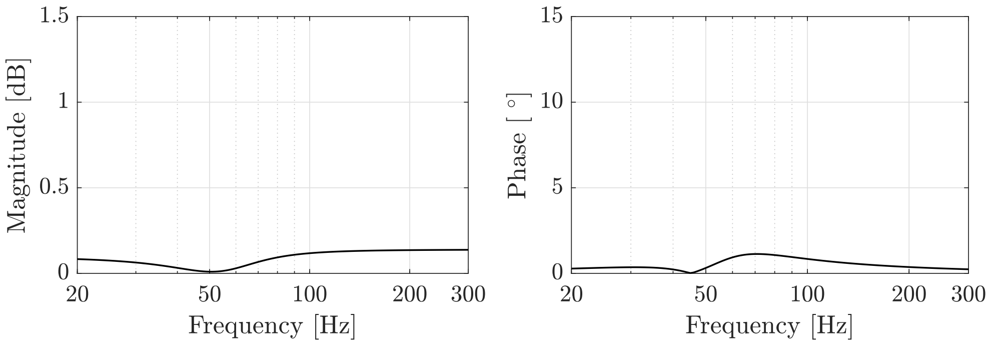
Research by Pedersen et al. [PA07] showed that the compliance C_{AS} can increase up to 25% when a speaker is running for 10 minutes. This was incorporated in the model assuming that the compliance increases by 25% for a voice coil temperature rise of 30 °C.

The resulting mean and standard deviation of the transfer functions used in the Monte-Carlo simulation are illustrated in figure 3.1b and 3.1c. All transfer functions were normalized such that the mean of the initial transfer function is 0 dB. The deviation is the smallest just below the closed-box resonance frequency at 57 Hz and increases both towards higher and lower frequencies. The greatest standard deviation is 0.14 dB at 300 Hz. Figure 3.1d shows the standard deviation of the phase in the MC simulation. The biggest phase errors occur just above the resonance frequency but never exceed 1.5°.



(a) Histogram of simulated voice coil temperatures

(b) Mean and standard deviation of normalized speaker transfer functions



(c) Standard deviation of speaker transfer function magnitude

(d) Standard deviation of speaker transfer function phase

Figure 3.1: Statistical properties of the loudspeaker transfer function

3.3 Modelling the room transfer function

Following sections 2.1.2 and 2.3.2 the room transfer functions were modeled by equation 2.17. According to ISO 7730:2006 [Sta15] moderate room temperatures are 20-24 °C in winter and 23-26 °C in summer. The speed of sound therefore varies from $c_{20^\circ} = 343\ \frac{\mathrm{m}}{\mathrm{s}}$ to $c_{26^\circ} = 347\ \frac{\mathrm{m}}{\mathrm{s}}$. It was assumed that all the temperatures are equally likely, so the modeled temperature follows a uniform distribution between 20 °C and 26 °C. The histogram of the temperatures used in the MC simulation are shown in figure 3.2a. Further it was assumed that the temperature varies uniformly across the room, i.e. there is no temperature gradient to the walls. The initial transfer function simulation was performed at 23 °C.

The mean transfer function from one speaker to one microphone at 200 different temperatures is shown in figure 3.2b with its standard deviation. A change in temperature stretches or squeezes the transfer function across the frequency axis as described in equation 2.46. The changes in the transfer function are more drastic at higher frequencies, as the modal density increases. Figure 3.2c illustrates this by showing the mean of all $M \times L$ standard deviations across the 200 MC iterations. A clear trend can be observed showing the greater deviations at higher frequencies. The same holds for the phase errors, which clearly increase with frequency.

3.4 Regularization strategies

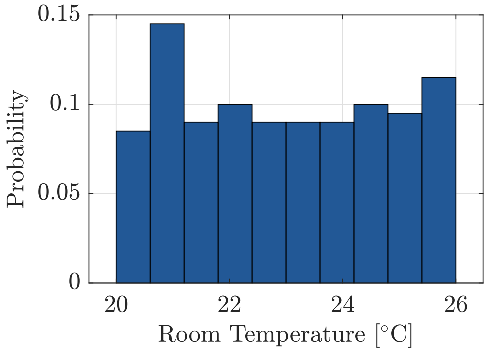
The influence of different regularization strategies is evaluated by means of the achieved acoustic contrast. The initial set of filters $\boldsymbol{w}_0$ is calculated based on the initial transfer function matrix $\boldsymbol{H}_0$ using equation 2.27. The achieved acoustic contrast from the initial transfer function and the matching filters is used as reference, since it corresponds to the optimal pressure matching solution. The effect that errors in the speaker and room transfer functions have on the acoustic contrast is analyzed in three different conditions: The transfer functions $\boldsymbol{H}_S$ contain errors in the speaker transfer functions, the transfer functions $\boldsymbol{H}_R$ contain errors in the room transfer functions and $\boldsymbol{H}_{SR}$ is the combination of both. The errors were modeled as described in section 3.2.2 and 3.3. The achieved acoustic contrast is averaged across the 200 MC iterations and the corresponding standard deviation is calculated.

3.4.1 Ridge regression

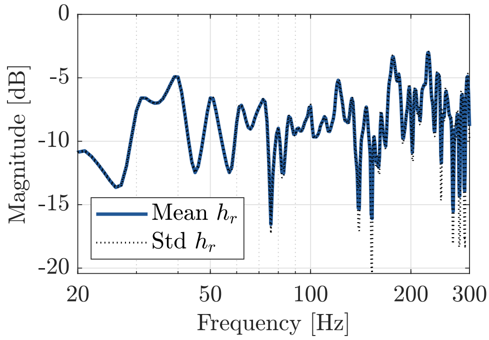
The effect of ridge regression on the robustness is evaluated on speaker and room transfer function errors. The filters are calculated for a set of regularization parameters λ from $\lambda = 10^{-5}$ to $\lambda = 10^{2}$. In this simulation the performance was evaluated for each λ and the frequency dependent λ_{opt} which showed the optimal performance was used for the MC simulation. This is an approach that is not feasible for feed-forward systems, so it could not be used in situ. Zhu et al. [Zhu+17] introduced regularization parameters based on the singular values, but at free field conditions. The reflections from the wall make the system far more complex, so finding a good regularization parameter can be a great challenge in practice [Zhu+17]. This report only shows the maximal robustness that can be gained from ridge regression by using the optimal regularization parameter.

3.4.2 Total least squares

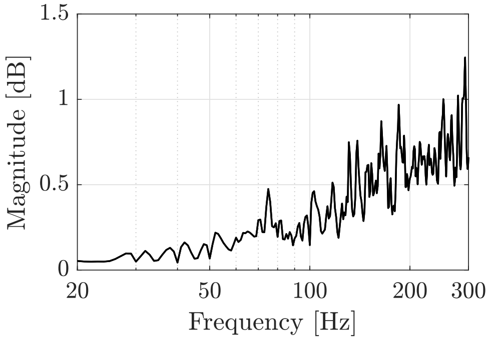
The TLS filters were determined from equation 2.44. Furthermore it was analysed whether regularized TLS can help making the system more robust. The optimal regularization parameter λ_{opt} was selected as in the RR simulation. The range of λ was set between 0 and the smallest singular value $\bar{\sigma}^2_{L+1}$. For $\lambda = \bar{\sigma}^2_{L+1}$ the RTLS equals the unregularized PM solution and for $\lambda = 0$ the RTLS equals the TLS.



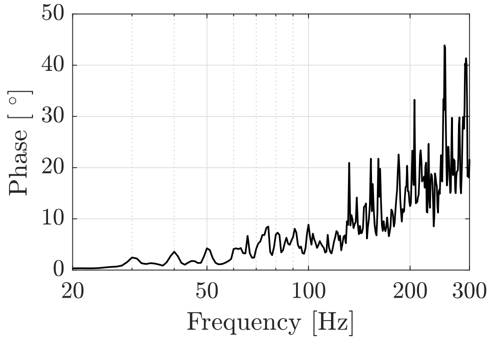
(a) Histogram of room temperature



(b) Mean and standard deviation of *one* room transfer function



(c) Mean of *all* $M \times L$ magnitude standard deviations across the 200 MC iterations



(d) Mean of *all* $M \times L$ phase standard deviations across the 200 MC iterations

Figure 3.2: Statistical properties of the room transfer functions

3.4.3 Resampling Temperature

The effect of resampling is simulated for three different filter implementations in the condition of varying room temperatures $\boldsymbol{H}_R$. The first two approaches, $\boldsymbol{w}_{0,res}$ and $\boldsymbol{w}_{0,res,err}$, are based on the initial filters $\boldsymbol{w}_0$ and resampled according to equation 2.47. For the resampling of the filters $\boldsymbol{w}_{0,res}$ the exact temperatures T_{new} and T_{init} match the real room temperatures that were used for the calculation of the room transfer functions. In situ these temperatures could never be determined with perfect accuracy. To simulate errors in the temperature measurement the filters $\boldsymbol{w}_{0,res,err}$ are based on normally distributed T_{new}, with a mean of the real temperature and a standard deviation of 1 °C. The third implemented filter $\boldsymbol{w}_{RR,res}$ is based on the RR filters without the presence of measurement errors.

The resampling of the filters is based on the *resample* function in *MATLAB* using linear interpolation. Since the filters are discrete resampling is an approximation that contains small errors.

3.4.4 Probability Model Optimization (PMO)

In order to calculate the PMO filters from equation 2.48 one needs to have good knowledge about the distribution of the underlying transfer function matrix. In this simulation the speaker and room transfer functions were modeled just as described in section 3.2 and 3.3, so the statistical properties of the distributions are the same as in the MC simulation. The calculation of the filters is based on 150 transfer function matrix simulations. The random seed was changed compared to the MC simulation such that the filters are based on transfer functions of the same statistical properties, but not on the exact same transfer functions as they are used on. The mean of these 150 matrices as well as the mean of the gram matrices was then utilized in equation 2.48 to gain the filters $\boldsymbol{w}_{PMO,S}$, $\boldsymbol{w}_{PMO,R}$ and $\boldsymbol{w}_{PMO,SR}$. Each of these filters was calculated by the same procedure for speaker errors, room transfer function errors and the combination of both.

CHAPTER 4

Simulation Results

In the following chapter the results from the simulations are presented. Goal of the simulations is to investigate how well different filters perform in different error conditions. Two different error conditions and their combination are simulated: Room transfer function errors, speaker transfer function errors and the combination of both. As described in chapter 3 they are simulated by a set of 200 transfer function matrices $\boldsymbol{H}_R$, $\boldsymbol{H}_S$ and $\boldsymbol{H}_{SR}$ respectively. The analyzed filters $\boldsymbol{w}_0$, representing the OLS filters, the RR filters $\boldsymbol{w}_{RR}$ and TLS filters $\boldsymbol{w}_{TLS}$ are based on the initial transfer function matrix $\boldsymbol{H}_0$. The filters $\boldsymbol{w}_{0,res}$, $\boldsymbol{w}_{0,res,err}$ and $\boldsymbol{w}_{RR,res}$ are resampled from $\boldsymbol{w}_0$ and $\boldsymbol{w}_{RR}$. The filters $\boldsymbol{w}_{PMO,S}$, $\boldsymbol{w}_{PMO,R}$ and $\boldsymbol{w}_{PMO,SR}$ are based on the probability distributions of the transfer function matrices in the presence of speaker transfer function errors, room transfer function errors and both. The results are summarized in table 4.1.

4.1 Influence of transfer function errors

The influence of speaker and room transfer function errors is illustrated in figure 4.1 and summarized in the first four rows of table 4.1. The results are compared to the achieved contrast when the initial filters $\boldsymbol{w_0}$ match the transfer function matrix $\boldsymbol{H}_0$. The result from the matching filter - transfer function combination are illustrated by the black line in figure 4.1. It can be observed that the optimal pressure matching filters achieve a high acoustic contrast of around 30 dB for most of the frequency range up to 100 Hz. For the higher frequencies the acoustic contrast decreases, such that the average acoustic contrast between 150 and 300 Hz is only 18.5 dB and the minimal acoustic contrast is only 8 dB at 294 Hz. Møller et al. [Møl+19] explain the poor acoustic contrast at higher frequencies by observing linear independence in the transfer function matrix. At high frequencies the modal density is high and the wavelengths are short, so the sound field changes rapidly across the room. Therefore the microphones sampling the sound field are fairly independent. Since there are only $L = 8$ loudspeakers to control the sound field at $M = 150$ microphone positions the system is overdetermined and the target sound field $\boldsymbol{p_t}$ is not in the span of $\boldsymbol{H_0}$. The resulting projection on to the span of $\boldsymbol{H_0}$ (the PM solution) is poorer, the reproduction error greater and the acoustic contrast smaller. At lower frequencies the wavelength increases and the modal density is lower. The sound field changes slowly across the room and adjacent microphones will be highly correlated. The system is less overdetermined and the PM solution has a smaller error, the acoustic contrast is greater.

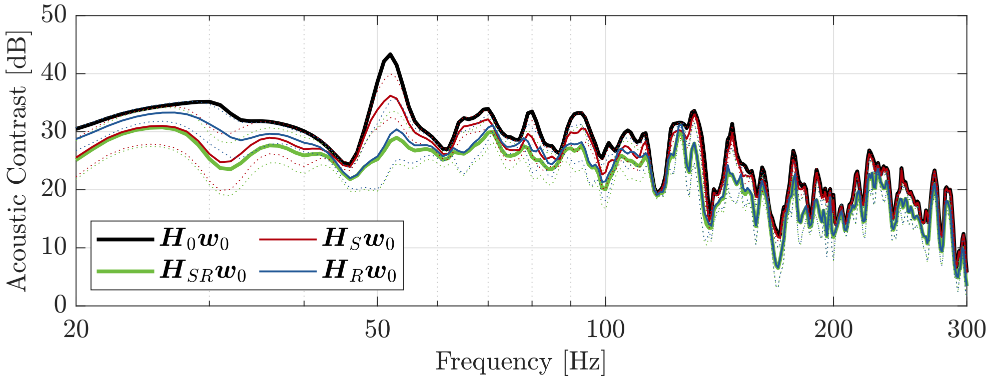


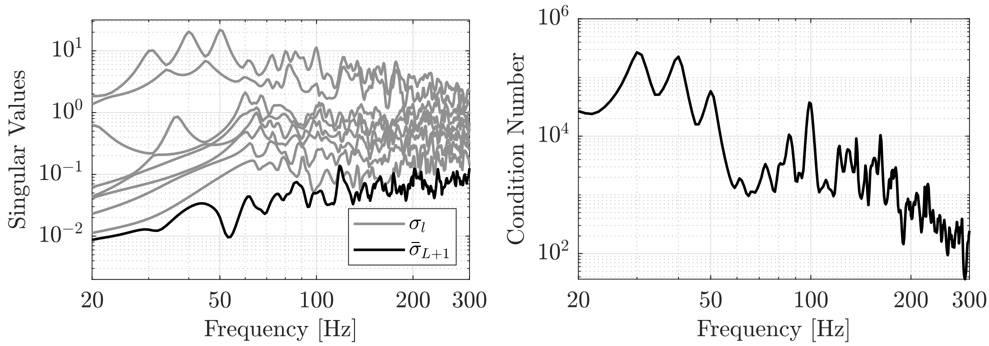
Figure 4.1: Influence of transfer function errors on the PM performance. Continous line: Mean AC across MC iterations, Dotted: ± Standard deviation across MC iterations

The achieved acoustic contrast using the initial filters $\boldsymbol{w}_0$ in the presence of speaker errors $\boldsymbol{H}_S$ is illustrated by the red line. The acoustic contrast is decreased by 3.5 dB in the low frequency range up to 60 Hz, but only by 0.4 dB in the high frequency range between 150 and 300 Hz. Looking at the standard deviations of the speaker phase and magnitude in figure 3.1c and 3.1d one can notice that the phase and magnitude errors are generally smaller at low frequencies. To explain why small errors result in greater decrease of contrast at low frequencies one can look at the condition number of the gram matrix $\boldsymbol{H}_0{}^H\boldsymbol{H}_0$ in figure 4.2b. As explained in section 2.4.2 a high condition number often implies a sensitive PM solution, as the individual columns of the transfer function matrix are correlated. The high condition number at low frequencies is a consequence of the high correlation between adjacent microphones. This high correlation in the row space makes the transfer function matrix nearly rank-deficient [Møl+19], so the first singular vector contains most of the information. The first singular value is high and the last singular value small, resulting in a great condition number. A consequence of the high condition number is that the PM solution can become very large. This can be expressed in terms of the control effort which measures the total signal power emitted to the room relative to a reference power. The control effort of the initial filters is plotted as the black line in figure 4.3. It can be observed that the control effort at very low frequencies below 50 Hz is significantly higher than at high frequencies. So the PM solution is very sensitive to transfer function errors at low frequencies and fairly robust at higher frequencies.

The effect of the correlation in the sampled sound field is therefore two-sided. The low correlation at high frequencies makes the system overdetermined resulting in a poorer, but more robust solution. At very low frequencies the system is almost underdetermined, good acoustic contrast can be achieved but at the cost of high control effort and sensitivity.

The blue line in figure 4.1 shows the effect of varying room temperatures on the acoustic contrast. At low frequencies below 45 Hz the impact of changing room temperatures is smaller than the impact of changing speaker transfer functions even though the standard deviations of magnitude and phase are greater (see figure 3.2c and 3.2d). The errors in the room transfer functions increase with frequency reaching up to 1 dB magnitude and 30° phase standard deviation. However due to the higher robustness at higher frequencies the impact in the frequency range up to 60 Hz is 3.8 dB but only 2.9 dB in the range from 150-300 Hz (see table 4.1).

The performance of the sound zones system with unregulated PM filters $\boldsymbol{w}_0$ in the presence of speaker and room transfer function errors is illustrated by the green line. Again the solution is the most sensitive at low frequencies such that the acoustic contrast is decreased by 5.7 dB in the low frequency range 20-60 Hz and 3.1 dB in the high frequency range 150-300 Hz.



(a) Singular values σ_l of transfer function matrix $\boldsymbol{H}_0$ and lowest singular value $\bar{\sigma}_{L+1}$ of the matrix $[\boldsymbol{H}_0,\ \boldsymbol{p}_t]$

(b) Condition number of ${\boldsymbol{H}_0}^H\boldsymbol{H}_0$

Figure 4.2: Singular values and condition number of initial transfer function matrix

4.2 Ridge regression

To show how big the impact of ridge regression can be on the performance of the system, the optimal regularization parameters λ_{opt} have to be determined first. Figure 4.4 shows how the regularization parameter was determined in the presence of speaker and room transfer function errors $\boldsymbol{H}_{SR}$. The mean reproduction error across the 200 MC iterations was calculated for each frequency and regularization parameter. The optimal reproduction error is illustrated by the green line. The red lines indicate where the reproduction error is increased by 1 dB relative to the optimal performance. If the regularization parameter is too high the system becomes too constrained by the regularization and the reproduction error increases rapidly. The introduced bias

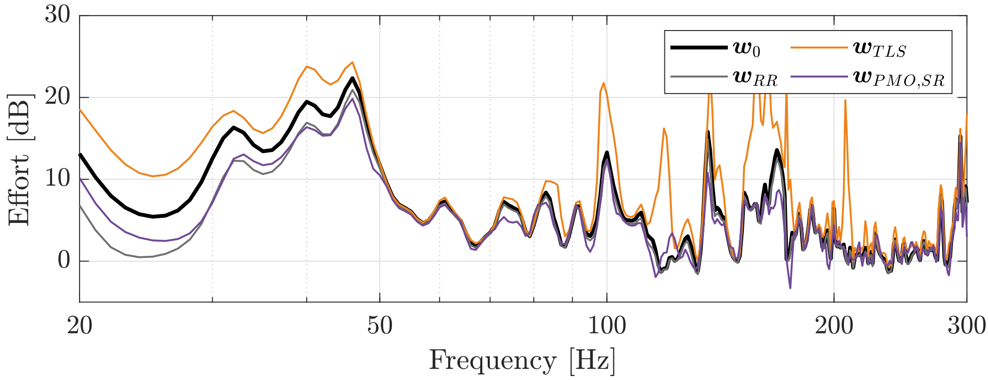


Figure 4.3: Control effort of different filters. Here the control effort is defined as the total signal power relative to the power required to achieve same sound pressure in the bright zone from the closest loudspeaker

makes the solution less sensitive to the underlying transfer functions and thereby fits the problem less accurately. If the regularization parameter is too low the solution equals the unregularized PM solution. It can be observed that the system is more sensitive to over- rather than underregularization. When choosing a too small λ the reproduction error is not increased by more than 1 dB for most of the frequency range. On the other hand choosing a λ greater than the bottom red line drastically reduces performance. A similar simulation was performed for the conditions $\boldsymbol{H}_S$ and $\boldsymbol{H}_R$ to determine the respective optimal regularization parameter.

The results from the ridge regression simulations in these three conditions are illustrated in figure 4.5 and summarized in table 4.1. The optimal filters $\boldsymbol{w}_{opt}$ are PM filters recalculated for each of the 200 iterations. The red, blue and green lines correspond to the respective conditions in figure 4.1.

Looking at figure 4.5a one can observe the influence of ridge regression on the robustness against speaker errors. Since the initial filters $\boldsymbol{w}_0$ are already quite robust to speaker errors at high frequencies the gain from ridge regression is small. At low frequencies however the mean acoustic contrast increased by 1.2 dB and the standard deviation dropped.

The influence of room transfer function errors in figure 4.5b is smaller at low frequencies and the ridge regression filters can achieve a similar acoustic contrast as the optimal filters up to 45 Hz. Between 45 Hz and 150 Hz the regularization does not have a significant impact. Above 150 Hz the acoustic contrast is increased by 0.6 dB relative to the unregularized filters.

In the presence of speaker and room transfer function errors the regularization has the biggest impact at low frequencies below 50 Hz. In the range between 50 and 150 Hz the influence is 0.5 dB. At the higher frequencies the regularization can gain 0.7 dB.

When comparing the OLS and RR solutions written in terms of the SVD in equations 2.34 and 2.40 it becomes apparent that RR reduces the sensitivity at low frequencies as the solution becomes less sensitive to errors in the direction of the low singular values. At higher frequencies the transfer function matrix is more orthogonal, the discrepancy between highest and smallest singular value decreases, errors in the direction of the small singular values have less impact on the filter norm and ridge regression is less effective. This can also be inferred from the control effort illustrated in figure 4.3. The high effort at low frequencies is significantly decreased by RR, but the impact decreases with frequency.

This little example showed that Tikhonov regularization alone can not regularize the problem. Another important factor to consider here is that in this case the optimal regularization parameter could be determined from multiple simulations. Usually the sound zones problem is designed as a feed-forward problem, making an exact parameter estimation difficult. Choosing a too high regularization parameter can significantly degrade performance making the small gain observed at low frequencies vanish easily.

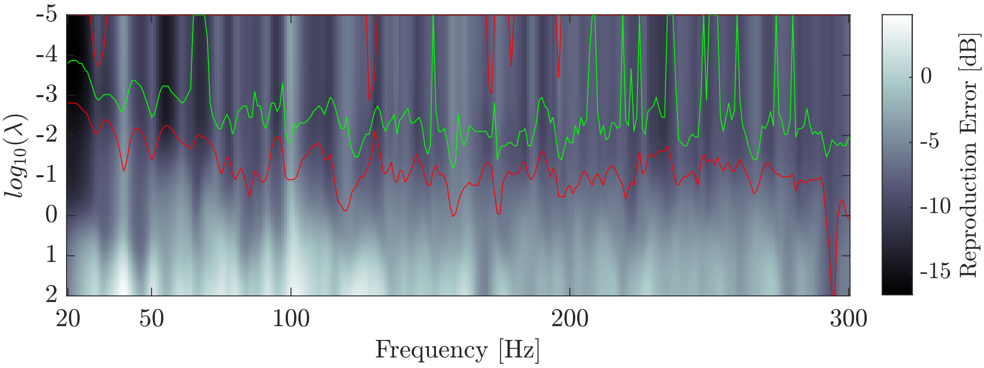
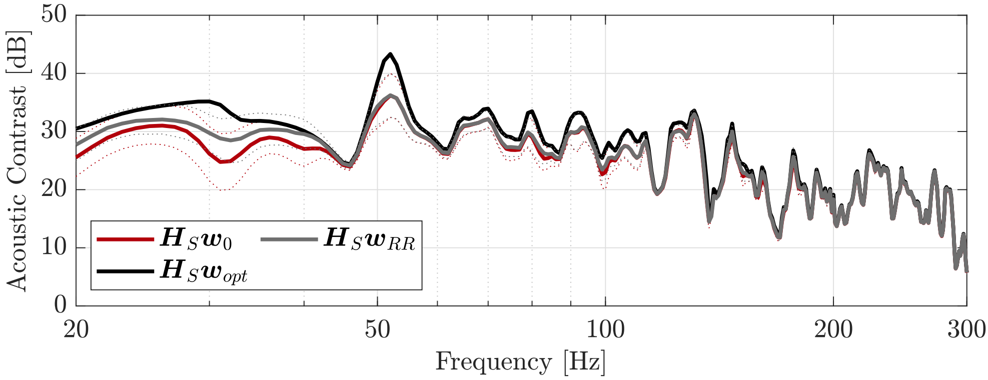


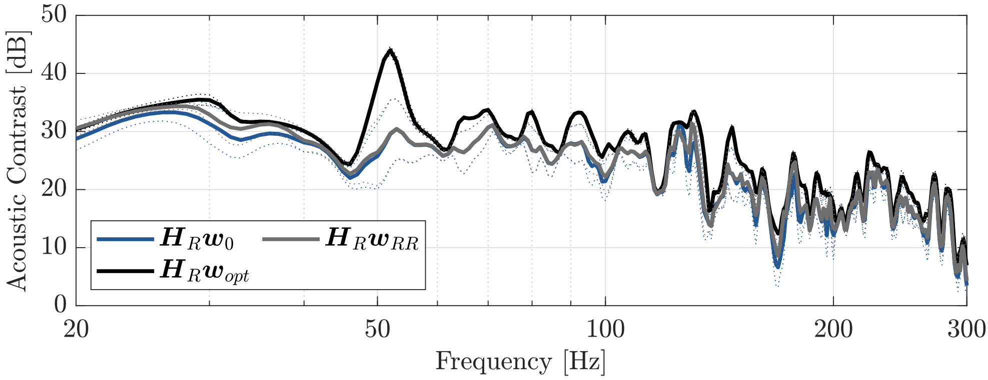
Figure 4.4: Reproduction Error as function of frequency and regularization parameter in presence of speaker and room transfer function errors $\boldsymbol{H}_{SR}$. The green line indicates λ_{opt}, the optimal RR regularization parameter yielding the lowest reproduction error. The red lines show where the system is underregularized (top red line) and overregularized (bottom red line) such that the reproduction error is increased by 1 dB relative to the optimal regularization. Note that the frequency axis is linear for better readability

4.3 Total Least Squares

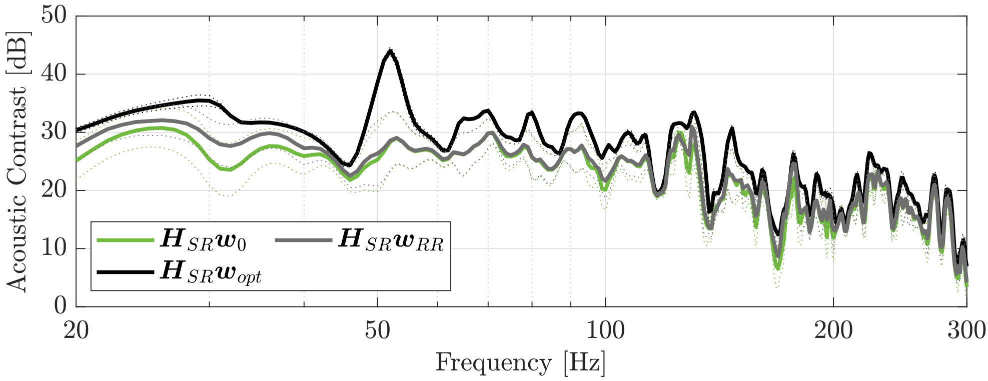
The Total Least Squares filter $\boldsymbol{w}_{TLS}$ are calculated using the smallest singular value $\bar{\sigma}_{L+1}$ of the matrix $[\boldsymbol{H}_0,\ \boldsymbol{p}_t]$. The smallest singular value $\bar{\sigma}_{L+1}$ is illustrated in figure 4.2a along to the singular values σ_l of the transfer function matrix $\boldsymbol{H}_0$. It can be



(a) Performance of ridge regression in presence of speaker transfer function errors $\boldsymbol{H}_S$



(b) Performance of ridge regression in presence of room transfer function errors $\boldsymbol{H}_R$



(c) Performance of ridge regression in presence of room and speaker transfer function errors $\boldsymbol{H}_{SR}$

Figure 4.5: Performance of ridge regression in different conditions. Continous line: Mean AC across MC iterations, Dotted: ± Standard deviation across MC iterations.

seen that $\bar{\sigma}_{L+1}$ is strictly smaller than the smallest σ_l, which can be shown to be the case for all TLS systems [The15]. It follows that the amount of deregularization $\bar{\sigma}^2_{L+1}$ increases with frequency.

The impact of TLS on the acoustic contrast is illustrated by the orange line in figure 4.6b. It can be observed that the deregularizing property of TLS decreases performance across the whole frequency range. The discrepancy is most severe at low frequencies, even though $\bar{\sigma}^2_{L+1}$ is the smallest in this frequency range. But since the deregularization increases the condition number of the inverted matrix it makes the TLS very sensitive to errors. This is further illustrated by the control effort in figure4.3. The effort of TLS filters is significantly higher, especially at low frequencies.

The selection of the optimal regularization parameter λ_{opt} is illustrated in figure 4.6a. The optimal RTLS regularization parameter is $\bar{\sigma}^2_{L+1}$ in almost the entire frequency range. At no point does RTLS decrease the reproduction error by more than 1 dB relative to the OLS solution.

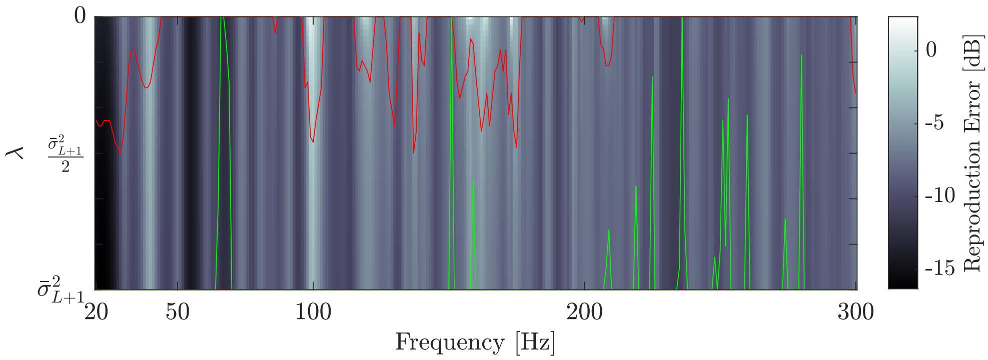
Overall TLS can not be used to make the PM system more robust to transfer function errors. The opposite is the case. TLS can be effective in the presence of small noise, but since the transfer function matrix can change significantly and TLS decreases the stability of the system it can not be used in this scenario.

4.4 Resampled Temperature

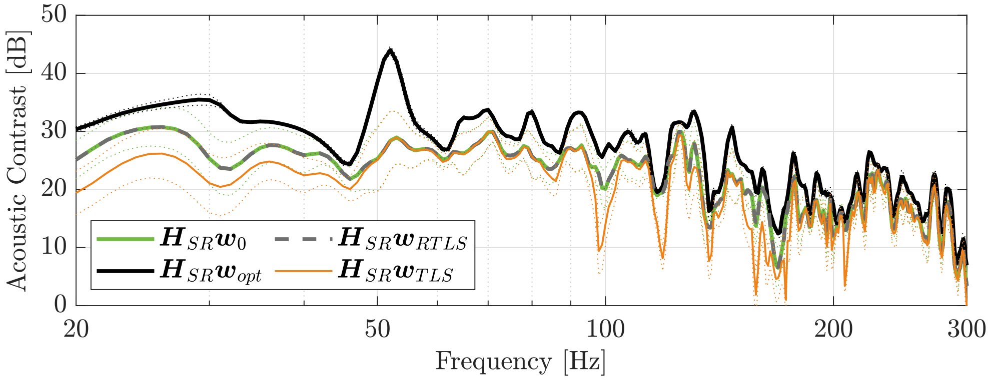
The performance of the resampled filter $\boldsymbol{w}_{0,res}$ and $\boldsymbol{w}_{0,res,err}$ in the presence of room transfer function errors $\boldsymbol{H}_R$ is illustrated in figure 4.7 and summarized in table 4.1. The performance of the resampled RR filter $\boldsymbol{w}_{RR,res}$ is not illustrated but can also be found in table 4.1.

It can be observed that resampling has high potential to mitigate the effect of room transfer function errors. Above the very low frequencies below 50 Hz the resampled filters $\boldsymbol{w}_{0,res}$ come very close to the optimal solution $\boldsymbol{w}_{opt}$. At low frequencies however the small inaccuracies caused by the resampling are enough to decrease the acoustic contrast by 1.4 dB in the range from 20 to 60 Hz. To come even closer to the optimal filters the resampled filters can be based on the more robust ridge regression filters. As shown in table 4.1 the filters $\boldsymbol{w}_{RR,res}$ achieve a better contrast across the whole frequency range than $\boldsymbol{w}_{0,res}$.

The performance of the filters simulating resampling on flawed temperature measurements $\boldsymbol{w}_{0,res,err}$ is illustrated by the teal line in figure 4.7. It shows that even when the temperature measurements are not exact they can help regularizing the problem. Across the whole frequency range these filters perform almost 2 dB better than the initial filters $\boldsymbol{w}_0$.



(a) Reproduction Error as function of frequency and regularization parameter in presence of speaker and room transfer function errors $\boldsymbol{H}_{SR}$. The green line indicates λ_{opt}, the optimal RTLS regularization parameter yielding the lowest reproduction error. The red line shows where the reproduction error is increased by 1 dB relative to the optimal regularization. The frequency axis is linear for better readability



(b) TLS and RTLS performance in presence of speaker and room transfer function errors $\boldsymbol{H}_{SR}$. The dotted lines indicate the ± standard deviations across MC iterations. The performance of RTLS is illustrated by a dashed line, as it equals the performance of the initial filter.

Figure 4.6: Performance of Total Least Squares and Regularized Total Least Squares in presence of speaker and room transfer function errors $\boldsymbol{H}_{SR}$

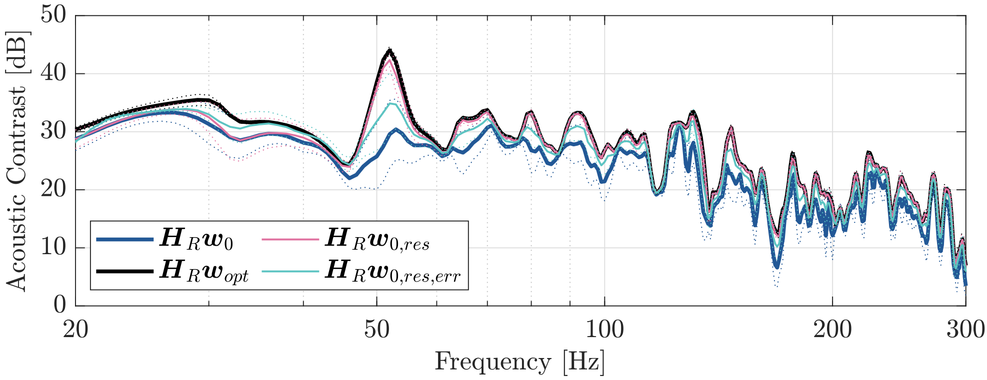


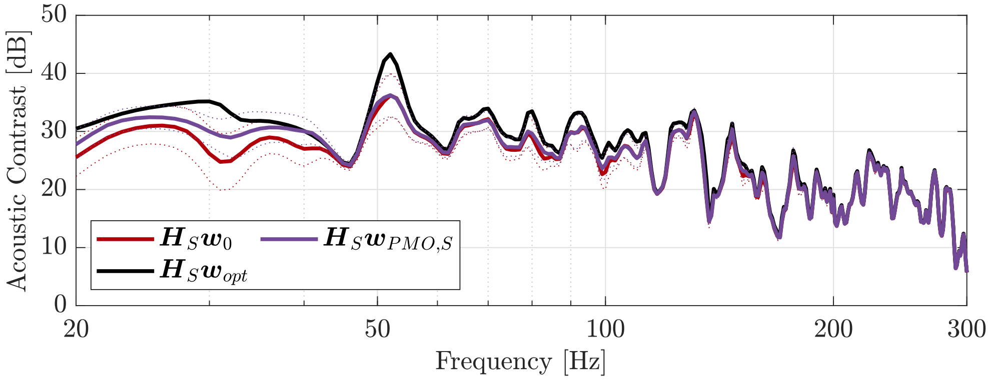
Figure 4.7: Performance of resampled filters $\boldsymbol{w}_{0,res}$ and $\boldsymbol{w}_{0,res,err}$ in presence of room transfer function errors $\boldsymbol{H}_R$ compared to optimal filters $\boldsymbol{w}_{opt}$ and initial filters $\boldsymbol{w}_0$

4.5 Probabiblity Model Optimization

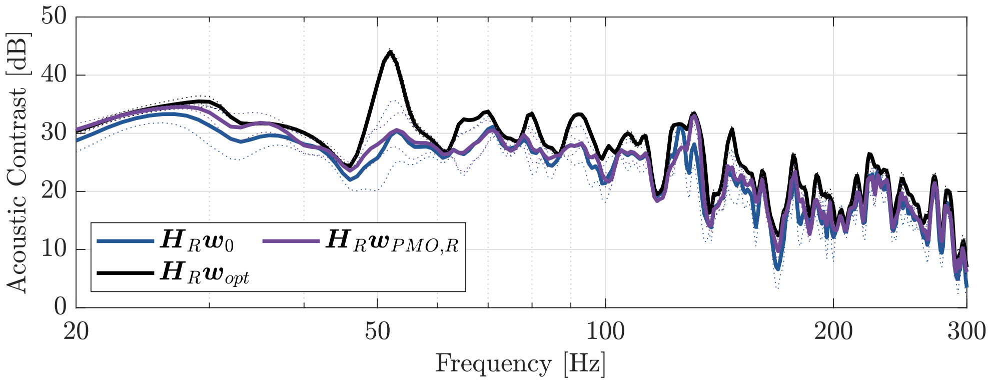
The results from PMO in the three error conditions is illustrated in figure 4.8. Figure 4.8a shows the robustness of the filters $\boldsymbol{w}_{PMO,S}$ against speaker errors. It can be observed that the regularizing effect above 60 Hz is marginal. Below 60 Hz however the mean acoustic contrast increased by 1.4 dB and the standard deviation is reduced.

The robustness against room transfer function errors is shown in figure 4.8b. Across the whole frequency range the acoustic contrast is increased and the standard deviation reduced relative to the initial filters $\boldsymbol{w}_0$. The regularization has the most significant impact at low frequencies. Below 60 Hz the mean acoustic contrast is increased by 1.3 dB. At higher frequencies the mean acoustic contrast increases by 0.4 db between 60 and 150 Hz and by 0.8 dB above 150 Hz. The standard deviation is reduced by $\sim$1 dB.

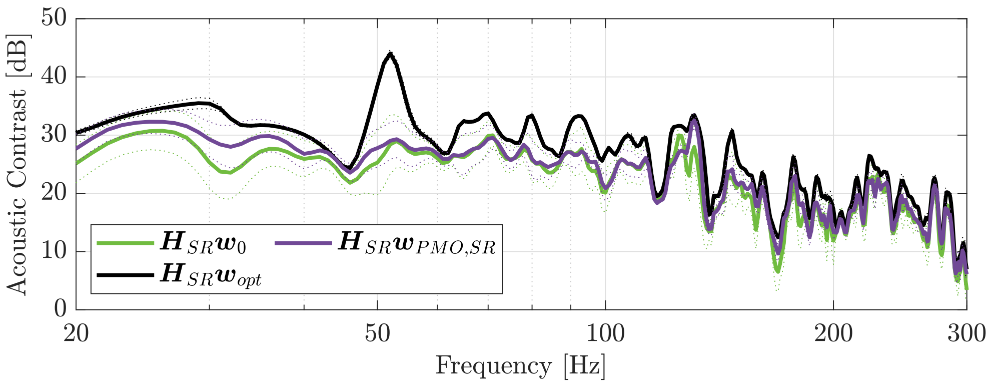
The last condition $\boldsymbol{H}_{SR}$ includes speaker and room transfer function errors. The filters $\boldsymbol{w}_{PMO,SR}$ have a big impact at low frequencies below 60 Hz, the mean acoustic contrast is raised from 26.6 dB to 28.5 dB and the results are a bit more stable with a standard deviation of 2.5 dB. At the frequencies above 60 Hz the mean acoustic contrast is only increased by $\sim$0.5 dB, but the standard deviation was reduced significantly by 0.9 dB.



(a) Performance of PMO in presence of speaker transfer function errors $\boldsymbol{H}_S$



(b) Performance of PMO in presence of room transfer function errors $\boldsymbol{H}_R$



(c) Performance of PMO in presence of room and speaker transfer function errors $\boldsymbol{H}_{SR}$

Figure 4.8: Performance of PMO in different conditions. Continous line: Mean AC, Dotted: $\pm$ Standard deviation

Condition	Filter	Mean AC [dB] 20-60 Hz	60-150 Hz	150-300 Hz
$\boldsymbol{H}_0$	$\boldsymbol{w}_0$	32.4 ± 0.0	28.3 ± 0.0	18.5 ± 0.0
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_0$	26.6 ± 3.3	24.0 ± 2.6	15.5 ± 2.2
$\boldsymbol{H}_S$	$\boldsymbol{w}_0$	28.9 ± 2.7	26.5 ± 1.6	18.2 ± 0.5
$\boldsymbol{H}_R$	$\boldsymbol{w}_0$	28.6 ± 2.5	24.8 ± 2.3	15.6 ± 2.1
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_{RR}$	28.1 ± 2.9	24.4 ± 2.3	16.1 ± 1.9
$\boldsymbol{H}_S$	$\boldsymbol{w}_{RR}$	30.1 ± 2.3	26.7 ± 1.4	18.2 ± 0.4
$\boldsymbol{H}_R$	$\boldsymbol{w}_{RR}$	29.5 ± 2.2	25.1 ± 2.1	16.2 ± 1.9
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_{TLS}$	24.0 ± 3.5	22.1 ± 2.9	13.7 ± 2.4
$\boldsymbol{H}_S$	$\boldsymbol{w}_{TLS}$	26.6 ± 3.0	25.0 ± 1.9	16.7 ± 0.7
$\boldsymbol{H}_R$	$\boldsymbol{w}_{TLS}$	26.1 ± 3.0	23.1 ± 2.5	13.9 ± 2.4
$\boldsymbol{H}_R$	$\boldsymbol{w}_{0,res}$	31.0 ± 1.6	28.0 ± 0.6	18.4 ± 0.9
$\boldsymbol{H}_R$	$\boldsymbol{w}_{0,res,err}$	30.6 ± 1.8	26.8 ± 1.6	17.4 ± 1.8
$\boldsymbol{H}_R$	$\boldsymbol{w}_{RR,res}$	31.8 ± 1.8	28.0 ± 1.6	18.5 ± 1.8
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_{PMO,SR}$	28.5 ± 2.5	24.5 ± 1.7	16.2 ± 1.3
$\boldsymbol{H}_S$	$\boldsymbol{w}_{PMO,S}$	30.3 ± 2.1	26.7 ± 1.4	18.2 ± 0.4
$\boldsymbol{H}_R$	$\boldsymbol{w}_{PMO,R}$	29.9 ± 1.6	25.2 ± 1.3	16.4 ± 1.2

Table 4.1: Table of simulation results. Mean acoustic contrast $\pm$ mean standard deviation. The mean standard deviation here is the standard deviation across MC iterations, averaged across the considered frequency bins. The mean AC is averaged across MC iterations and frequency bins.

CHAPTER 5

Discussion

The purpose of the performed simulations was to investigate two research questions. First, it was asked how big the impact of transfer function errors is on the performance of the sound zones system. The second question asked, to which extent different regularization strategies can reduce the sensitivity of the system. The following chapter discusses the results and limitations of the performed simulations with respect to the proposed research questions.

5.1 Limitations of Simulations

The impact of transfer function errors on the performance of the sound zones system was investigated by simulating speaker and room transfer function errors. Consequently the quality of the simulation results is determined by the validity of the transfer function simulations.

The loudspeaker transfer function model, as described in section 3.2, focuses on changes in the DC-Resistance of the coil and changes in suspension compliance. Non-linear effects such as distortion products or level-dependent transfer function changes are not considered. The impact of non-linear distortion is assumed to be small in the considered frequency range [Ma+19], but level-dependent transfer function changes can not be ruled out [Ma+18]. The introduced error model provides a framework for error simulations and an estimate for the magnitude of the assumed loudspeaker transfer function changes, however reliable measurements of a speaker transfer function are necessary to validate this estimation.

The room transfer functions underlie a major simplification: It is assumed that the temperature changes uniformly across the room. In situ this will never be the case, every room has temperature gradients. The radiator has a different temperature than the windows and walls, the ceiling will differ from the floor. To simulate the transfer functions in the room in the presence of temperature gradients requires a more elaborate numerical simulation since an analytical solution of the wave equation does not exist in this case. The performed simulations can therefore only give an indication of how big the influence of temperature changes is.

The second question of this report is, to which extent the impact of transfer function errors can be regularized using more advanced filter design. This was investigated using four different filter concepts: The ridge regression filters $\boldsymbol{w}_{RR}$, the TLS filters $\boldsymbol{w}_{TLS}$, the resampled filters $\boldsymbol{w}_{0,res}$ and the PMO filters $\boldsymbol{w}_{PMO}$. Besides the TLS filters, these different filters have in common that their performance relies on more information about the underlying system than just the initial transfer function matrix

$\boldsymbol{H}_0$. However, in situ this information might not be available, which could potentially degrade performance significantly.

For the RR simulation the optimal filters were determined by the feed-back from the simulations. The reproduction error was determined for a set of regularization parameters. Based on these simulations the optimal filters were determined. This sort of feed-back filter design is only available if the changes of the transfer function matrix are known. In situ the choice of regularization parameter can be challenging. Choosing a too high regularization parameter can deteriorate performance significantly.

The performance of the resampled filters depends on two parameters: How well the temperature can be measured in the room and how well the transfer functions can be resampled. In the performed simulations it was assumed that the temperature changes uniformly across the room, and that the boundary conditions of the room do not change. Therefore the transfer functions can be retrieved nearly perfectly when the temperature change is known. In situ this could not be assumed, affecting the expected performance of the filters in situ. Olsen et al. [OM17] investigated resampling in the sound zones problem in cars. It was shown that resampling was not sufficient to have robust performance. However, their temperature variations were much bigger ($T_{cold} = -2$ °C and $T_{hot} = 22.8$ °C) and the boundary conditions in a car are very different. It is assumed that the boundary conditions in cars, i.e. the absorbing properties of the leather seats and interior, are temperature dependent. In rooms most of the boundary conditions are defined by the walls, which might be less temperature dependent. It has to be investigated in the future how well resampling works in domestic rooms, when the temperature is non-uniform across the room.

The efficiency of the PMO filters is determined by the agreement of the statistical transfer function properties used for determining the filters, and the real statistical properties. In this simulation both distributions were generated by the same procedure, so their agreement is maximal. In situ an exact estimation of the transfer function errors is impossible, so the performance of the simulated PMO filters is better than possible in a real room.

Zhu et al. [Zhu+17] performed simulations and measurements analysing PMO filters. The setup considered in that paper is different, the filter performance is evaluated in an acoustically treated studio environment and the filters are based on acoustic contrast control. One aspect of that paper was applying uniformly distributed multiplicative errors (ME) with an error bound of ± 1 dB to the transfer functions and comparing the performance of different filters. Two of these filters were PMO filters, designed by the same principle, but one maximizes the acoustic contrast for an assumed uniform multiplicative error bound of ± 1 dB and one for an error bound of ± 3 dB. They showed that both filters achieve similar performance when evaluated on the the same multiplicative errors.

This motivated further research for this thesis, to investigate whether the PMO filters achieve similar contrast when the PMO filters are not based on the distributions they are evaluated on. It was investigated whether a precise error model is necessary to achieve good performance from PMO filters by evaluating the performance of less accurate PMO filters. The results are shown in table 5.1. The filters $\boldsymbol{w}_{PMO_{Over}}$ and $\boldsymbol{w}_{PMO_{Under}}$ are both based on the physical error model used to simulate the speaker and room transfer function errors. The filter $\boldsymbol{w}_{PMO_{Over}}$ however overestimates the errors in the system, by assuming a uniform room temperature distribution from 17-29 °C and a wider spread gamma distribution of the voice coil temperature. The filter $\boldsymbol{w}_{PMO_{Under}}$ underestimates the system errors by assuming a uniform room temperature distribution between 22 and 24 °C and a narrower voice coil temperature distribution. It is observed that the underestimating filter has a higher standard deviation than the overestimating filter and that the performance of both filters is worse than that of the matching PMO filters $\boldsymbol{w}_{PMO,SR}$. At low frequencies below 60 Hz however the sacrifice in mean acoustic contrast is small and the standard deviation is decreased significantly when overestimating the error sources. The filters $\boldsymbol{w}_{PMO_{ME0.1}}$, $\boldsymbol{w}_{PMO_{ME0.5}}$, $\boldsymbol{w}_{PMO_{ME1}}$ and $\boldsymbol{w}_{PMO_{ME3}}$ assume uniform multiplicative errors with error bounds of 0.1, 0.5, 1 and 3 dB, and phase error bounds of 0.6, 2.9, 5.7 and 17° to emulate the filters from Zhu et al. It can be observed that none of the multiplicative error PMO filters can compete with the matching filter $\boldsymbol{w}_{PMO,SR}$ across the whole frequency range and that the filters perform best, when they slightly overestimate the underlying error distributions (compare figures 3.1 and 3.2). The filter $\boldsymbol{w}_{PMO_{ME0.5}}$ performs well at low frequencies where the transfer function errors are small but significantly worse than the $\boldsymbol{w}_{PMO,SR}$ filter at high frequencies where the transfer function errors are higher. On the other hand the filter $\boldsymbol{w}_{PMO_{ME3}}$ performs well at high frequencies but poorly at low frequencies. To attain a good PMO filter it is therefore important to have a good error model, that allows to estimate errors across the whole frequency range.

Overall, the simulated filters $\boldsymbol{w}_{RR}$, $\boldsymbol{w}_{0,res}$ and $\boldsymbol{w}_{PMO}$ should be seen as ideal filters. So the simulations assess the possible gain in robustness provided precise knowledge is available about the error sources.

5.2 Discussion of results

5.2.1 Effect of transfer function errors

The first research question of this thesis asked how big the impact of transfer function errors is on the performance of the system. The results were analysed in section 4.1 and illustrated in figure 4.1. While the system shows the best performance at low frequencies, it is also most sensitive due to the high condition number. This can be seen by the influence of the speaker transfer function errors. The speaker transfer function errors are in a similar order of magnitude at high frequencies 150-300 Hz and at low frequencies 20-60 Hz (see figure 3.1c and 3.1d). However, the speaker errors hardly have an impact on the acoustic contrast performance at high frequencies, while the mean acoustic contrast is decreased by 3.5 dB at low frequencies. Further, since the standard deviation across MC iterations is much higher at low frequencies the maximal loss of acoustic contrast is even greater.

Condition	Filter	Mean AC [dB] 20-60 Hz	60-150 Hz	150-300 Hz
$\boldsymbol{H}_0$	$\boldsymbol{w}_0$	32.4 ± 0.0	28.3 ± 0.0	18.5 ± 0.0
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_0$	26.6 ± 3.3	24.0 ± 2.6	15.5 ± 2.2
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_{PMO,SR}$	28.5 ± 2.5	24.5 ± 1.7	16.2 ± 1.3
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_{PMO_{Over}}$	28.2 ± 1.8	23.6 ± 1.3	15.2 ± 0.9
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_{PMO_{Under}}$	27.9 ± 3	24.6 ± 2.3	15.9 ± 2
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_{PMO_{ME0.1}}$	26.9 ± 3.3	24 ± 2.6	15.5 ± 2.2
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_{PMO_{ME0.5}}$	28.3 ± 2.8	24.1 ± 2.6	15.5 ± 2.2
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_{PMO_{ME1}}$	28.2 ± 2.2	24.4 ± 2.4	15.6 ± 2.1
$\boldsymbol{H}_{SR}$	$\boldsymbol{w}_{PMO_{ME3}}$	25 ± 0.8	23.4 ± 1.8	16 ± 1.9

Table 5.1: Table of simulation results of different PMO implementations. Mean acoustic contrast ± mean standard deviation. Again, the mean standard deviation here is the standard deviation across MC iterations, averaged across the considered frequency bins. The mean AC is averaged across MC iterations and frequency bins. The top three condition - filter combinations are a reference from table 4.1. The filters $\boldsymbol{w}_{PMO_{Over}}$ and $\boldsymbol{w}_{PMO_{Under}}$ are determined from an over- and underestimated version of the error model. The filters $\boldsymbol{w}_{PMO_{ME0.1}}$, $\boldsymbol{w}_{PMO_{ME0.5}}$, $\boldsymbol{w}_{PMO_{ME1}}$ and $\boldsymbol{w}_{PMO_{ME3}}$ assume uniform multiplicative errors with error bounds of 0.1, 0.5, 1 and 3 dB, and phase error bounds of 0.6, 2.9, 5.7 and 17° respectively. The error condition $\boldsymbol{H}_{SR}$ contains speaker and room transfer function errors. The condition $\boldsymbol{H}_0$ is the initial transfer function matrix on which the filter $\boldsymbol{w}_0$ is based.

At low frequencies the influences of speaker and room transfer function errors on the acoustic contrast are in a similar order of magnitude. At higher frequencies the room transfer function magnitude and phase errors increase (see figures 3.2c and 3.2d), such that the room transfer function errors also have a significant impact at higher frequencies.

The loss of acoustic contrast at higher frequencies is critical, as even the optimal filters only achieve an acoustic contrast of 18.5 dB (150-300 Hz). To attain a clearer picture of the loss in acoustic contrast, the mean sound pressure level in the two zones is illustrated in figure 5.1. This plot shows the impact of the combined speaker and room transfer function errors for the optimal filters $\boldsymbol{w}_{opt}$, initial PM filters $\boldsymbol{w}_0$, RR filters $\boldsymbol{w}_{RR}$ and PMO filters $\boldsymbol{w}_{PMO,SR}$. The SPL in the bright zone is almost unaffected by the transfer function error. The dark zone SPL, however, is increased significantly in the presence of transfer function errors. This discrepancy can be explained by the logarithmic scaling of the SPL. While the reproduction error is in the same order of magnitude in both zones, the effect on the SPL is enhanced at low levels. It is therefore hard to maintain silence in the dark zone.

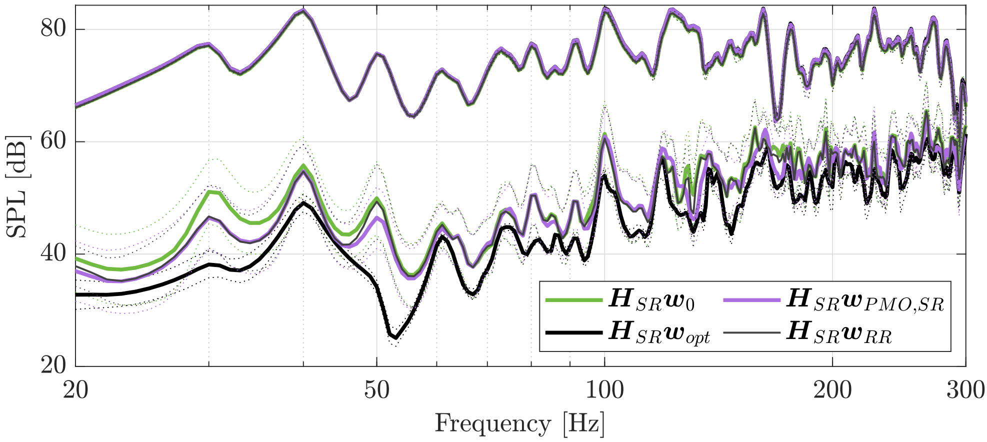


Figure 5.1: Mean Sound Pressure Level in bright (top lines) and dark zone (bottom lines) in presence of speaker and room transfer function errors for different filters. Dotted lines indicate ± standard deviation across MC iterations.

5.2.2 Comparison of different filters

Out of all the analysed filters the Total Least Squares filters $\boldsymbol{w}_{TLS}$ show the worst performance. The introduced deregularization by the squared smallest singular value $\bar{\sigma}^2_{L+1}$ destabilizes the system. The mean acoustic contrast is significantly decreased compared to the initial filters $\boldsymbol{w}_0$ in all error conditions. TLS is therefore not suitable for the given sound zones setup.

Assuming that the temperature in the room changes uniformly, and that the boundary conditions of the room are not temperature dependent, resampling can be very effective at retrieving acoustic contrast. By also simulating measurement errors with the filter $\boldsymbol{w}_{0,res,err}$ it was verified that the acoustic contrast performance can be improved with resampling even when the measurements are not slightly inaccurate. Especially at high frequencies resampling was shown to be effective. The performance at low frequencies can further be improved by resampling the RR filters $\boldsymbol{w}_{RR}$.

Ridge regression is most effective when the sensitivity is caused by collinearity in the transfer function matrix. At low frequencies the sound field sampled by the microphones is highly correlated [Møl+19], so the control effort is high and ridge regression is effective. At higher frequencies the regularizing effect is less prominent. Since over-regularizing can deteriorate system performance drastically it is safer to choose too little regularization when establishing a feed-forward system. This however reduces the achievable gain, making the effect of high frequency regularization negligible.

The PMO filters increase robustness of the system to transfer function errors. Both, the mean and the standard deviation are improved compared to the initial filters $\boldsymbol{w}_0$. Especially at low frequencies the mean acoustic contrast can be increased significantly. At high frequencies the improvement in mean acoustic contrast is less prominent, but the standard deviation is significantly reduced in the presence of room transfer function changes.

Compared to the RR filters the PMO filters have a slightly higher mean acoustic contrast for all frequencies and error conditions. The standard deviation is significantly reduced in the presence of room transfer function errors.

Overall, the PMO filters achieve high acoustic contrast with a low standard deviation. However, as discussed earlier, the PMO filters were based on the same statistical model as the error model on which the performance was evaluated on. Basing the PMO filters on different error models decreases performance.

In general, when comparing the resampled filters, RR filters and PMO filters it becomes apparent that an increasing amount of knowledge about the underlying transfer function errors can be leveraged to improve the sound zones performance. An exact temperature measurement allows to resample the filters, which can almost retain optimal performance. To achieve this near optimal performance the filters have to be updated constantly, whenever the room temperature changes. The resampled filters thereby depend on a current transfer function estimation, whereas the RR and PMO filters can be formulated based on the statistical properties of the transfer function distributions. The PMO filters perform better than the RR filters, when the underlying distribution is known, however when the PMO filters are based on over- or underestimations of the transfer function errors this high performance can not be sustained. The RR filters on the other hand also depend on a good estimation of the regularization coefficient. A poor estimation of the regularization coefficient can significantly reduce performance. Therefore, the performance of sound zones systems can significantly be improved by accurate error models.

CHAPTER 6

Conclusion

In this thesis an error model was proposed that aims at simulating transfer function errors that can arise in the sound zones context at low frequencies. The focus thereby was on simulating room transfer function errors resulting from changes in room temperature and changes in the loudspeaker transfer function.

Further, the influence of these transfer function errors was investigated from Monte-Carlo simulations. It was shown that changes in room temperature can have a significant impact on the acoustic contrast across the whole frequency range from 20-300 Hz. The influence of the loudspeaker transfer function errors on the performance of the system was most prominent at low frequencies. At higher frequencies from 150-300 Hz the loudspeaker errors only decreased performance marginally.

Different pressure matching filters were examined in terms of their performance in the presence of the loudspeaker and room transfer function errors. The Total Least Squares filters reduced robustness to errors and decreased acoustic contrast, so the TLS filters are not suitable for the given scenario. While the initial ordinary least squares filters and the TLS filters can be determined from a single transfer function matrix, the other filters require further information about the system errors to be determined correctly. Under the assumption that the room temperature is uniform across the room, resampling can retrieve near optimal performance when the room temperature is determined exactly. For the design of ridge regression filters one needs to determine a regularization coefficient, that increases robustness to errors. However, to determine this coefficient optimally, the transfer function errors need to be estimated. A similar statement can be formulated for the PMO filters. While the PMO filters can achieve significant improvements in mean acoustic contrast and standard deviation, their performance depends on a precise error model. When choosing the wrong regularization coefficient for the RR filters, or basing the PMO filters on inaccurate error estimations can significantly impair performance. Therefore a well defined error model, that can estimate the impact of different error sources is essential for the design of robust pressure matching filters.

Future research could validate the performed simulations with measurements. It would be of great interest to find out how well the introduced error model can simulate in situ transfer function errors. If the model does a fair job at simulating the transfer function errors, PMO or RR filters could easily be designed for different sound zones setups that could increase robustness. Further, it would be interesting to investigate how well temperature dependent transfer function changes can be inferred from resampling in the presence of temperature gradients in the room. This could provide further insights into how effective resampling can be to be increase robustness to temperature dependent transfer function changes.

Bibliography

[Aba17] Ahmad Abawi. "Finite Element and Boundary Methods in Structural Acoustics and Vibration Finite Element and Boundary Methods in Structural Acoustics and Vibration , Atalla Noureddine Sgard Franck CRC Press, Taylor Francis Group, Boca Raton, London, New York, ISBN 13:978-1-4665-9287-2". In: *The Journal of the Acoustical Society of America* 141 (June 2017), pages 4300–4300. DOI: 10.1121/1.4984771.

[Age07] Finn Agerkvist. "modelling loudspeaker non-linearities". In: *journal of the audio engineering society* (September 2007).

[Bay+15] Khan Baykaner et al. "the relationship between target quality and interference in sound zone". In: *journal of the audio engineering society* 63.1/2 (January 2015), pages 78–89. DOI: https://doi.org/10.17743/jaes.2015.0007.

[Ben20] Colin Benker. "Simulation study on perturbations in the sound zones problem". In: (2020). URL: https://gitlab.gbar.dtu.dk/s190247/ResearchProject/blob/b495d6cd20b2e1c496faaac5510501d372f6e112/SimulationStudyOnPerturbationsInTheSoundZoneProblem.pdf.

[Cha98] Peter John Chapman. "thermal simulation of loudspeakers". In: *journal of the audio engineering society* (May 1998).

[CJ12] Jiho Chang and Finn Jacobsen. "The effect of scattering on sound field control with a circular double-layer array of loudspeakers". English. In: *Proceedings of the 132nd AES Convention.* AES 132nd Convention ; Conference date: 26-04-2012 Through 29-04-2012. AES, 2012, pages 707–715. ISBN: 9781622761180. URL: http://www.aes.org/.

[CJ13] Ji-Ho Chang and Finn Jacobsen. "Sound field control with a circular double-layer array of loudspeakers". In: *The Journal of the Acoustical Society of America* 133 (April 2013), pages 2046–54. DOI: 10.1121/1.4792486.

[CK02] Joung-Woo Choi and Yang-Hann Kim. "Generation of an acoustically bright zone with an illuminated region using multiple sources". In: *The Journal of the Acoustical Society of America* 111.4 (2002), pages 1695–1700. DOI: 10.1121/1.1456926. eprint: https://doi.org/10.1121/1.1456926. URL: https://doi.org/10.1121/1.145692.

[Col+14a] P. Coleman et al. "Personal audio with a planar bright zone." In: *The Journal of the Acoustical Society of America* 136 4 (2014), pages 1725–35.

[Col+14b] Philip Coleman et al. "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array". In: *The Journal of the Acoustical Society of America* 135.4 (2014), pages 1929–1940. DOI: 10.1121/1.4866442. eprint: https://doi.org/10.1121/1.4866442. URL: https://doi.org/10.1121/1.4866442.

[DG97] W. F. Druyvesteyn and John Garas. "Personal Sound". In: *J. Audio Eng. Soc* 45.9 (1997), pages 685–701. URL: http://www.aes.org/e-lib/browse.cfm?elib=7843.

[GL13] Gene H. Golub and Charles F. van Loan. *Matrix Computations*. Fourth. JHU Press, 2013. ISBN: 1421407949 9781421407944. URL: http://www.cs.cornell.edu/cv/GVL4/golubandvanloan.htm.

[Gui+15] P. Guidorzi et al. "Impulse Responses Measured with MLS or Swept-Sine Signals Applied to Architectural Acoustics: An In-depth Analysis of the Two Methods and Some Case Studies of Measurements Inside Theaters". In: *Energy Procedia* 78 (2015). 6th International Building Physics Conference, IBPC 2015, pages 1611–1616. ISSN: 1876-6102. DOI: https://doi.org/10.1016/j.egypro.2015.11.236. URL: http://www.sciencedirect.com/science/article/pii/S1876610215019682.

[HO96] Per Christian Hansen and Dianne P. O'Leary. "Regularization algorithms based on total least squares". eng. In: *Recent Advances in Total Least Squares Techniques and Errors-in-variables* (1996). Edited by Sabine Van Huffel, pages 127–137.

[Jac+11] Finn Jacobsen et al. "A comparison of two strategies for generating sound zones in a room." English. In: *Proceedings of 18th International Congress on Sound and Vibration*. 18th International Congress on Sound and Vibration, ICSV 18 ; Conference date: 10-07-2011 Through 14-07-2011. International Institute of Acoustics and Vibration, 2011. ISBN: 978-85-63243-01-0. URL: http://www.icsv18.org/.

[JJ13] Finn Jacobsen and Peter Møller Juhl. *Fundamentals of General Linear Acoustics*. English. United States: John Wiley Sons Ltd, July 2013. ISBN: 9781118346419.

[Kas05] Safa Kasap. *Principles of Electronic Materials and Devices*. 3rd edition. USA: McGraw-Hill, Inc., 2005. ISBN: 0073104647.

[KN93] Ole Kirkeby and Philip A. Nelson. "Reproduction of plane wave sound fields". In: *The Journal of the Acoustical Society of America* 94.5 (1993), pages 2992–3000. DOI: 10.1121/1.407330. eprint: https://doi.org/10.1121/1.407330. URL: https://doi.org/10.1121/1.407330.

[Lim17] Tymphany HK Limited. *Driver Specification Sheet SLS-P830668*. 2017. URL: https://www.tymphany.com/drivers/driver_functions/specsheet.php?id=814.

[LL03] W.M. Leach and M. Leach. *Introduction to Electroacoustics and Audio Amplifier Design*. Kendall/Hunt Publishing Company, 2003. URL: https://books.google.dk/books?id=Nx0fAQAAIAAJ.

[Ma+18] Xiaohui Ma et al. “Impact of loudspeaker nonlinear distortion on personal sound zones”. English. In: *Acoustical Society of America. Journal* 143.51 (January 2018), pages 51–59. ISSN: 0001-4966. DOI: 10.1121/1.5019476.

[Ma+19] Xiaohui Ma et al. “Nonlinear distortion reduction in sound zones by constraining individual loudspeaker control effort”. eng. In: *Journal of the Audio Engineering Society* 67.9 (2019), pages 641–654. ISSN: 15494950, 00047554. DOI: 10.17743/jaes.2019.0015.

[Møl+19] M. B. Møller et al. “On the Influence of Transfer Function Noise on Sound Zone Control in a Room”. In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 27.9 (2019), pages 1405–1418.

[OM17] Martin Olsen and Martin Bo Møller. “Sound zones: On the effect of ambient temperature variations in feed-forward systems”. English. In: *142nd Audio Engineering Society International Convention 2017*. 142nd Audio Engineering Society International Convention 2017, AES 2017 ; Conference date: 20-05-2017 Through 23-05-2017. Audio Engineering Society, 2017, pages 1009–1018. ISBN: 9781510843523.

[PA07] Bo Rohde Pedersen and Finn T. Agerkvist. “Time Varying Behavior of the Loudspeaker Suspension”. English. In: *Audio Engineering Society(AES) convention*. AES 123rd. convention ; Conference date: 01-01-2007. 2007.

[PCK13] Jin-Young Park, Jung-Woo Choi, and Yang-Hann Kim. “Acoustic contrast sensitivity to transfer function errors in the design of a personal audio system”. In: *The Journal of the Acoustical Society of America* 134.1 (2013), EL112–EL118. DOI: 10.1121/1.4809778. eprint: https://doi.org/10.1121/1.4809778. URL: https://doi.org/10.1121/1.4809778.

[PM06] John G. Proakis and Dimitris K Manolakis. *Digital Signal Processing (4th Edition)*. 4th edition. Prentice Hall, 2006. ISBN: 0131873741.

[Ros07] T. Rossing. *Springer Handbook of Acoustics*. Springer Handbook of Acoustics. Springer New York, 2007. ISBN: 9780387304465. URL: https://books.google.dk/books?id=4ktVwGe%5C_dSMC.

[SK01] Ulf Seidel and Wolfgang Klippel. “fast and accurate measurement of the linear transducer parameters”. In: *journal of the audio engineering society* (May 2001).

[Sta15] International Organization for Standardization. "DS/ISO 7730: Ergonomics of the thermal environment – Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria". In: (November 2015).

[The15] Sergios Theodoridis. "Chapter 6 - The Least-Squares Family". In: *Machine Learning*. Edited by Sergios Theodoridis. Oxford: Academic Press, 2015, pages 233–274. ISBN: 978-0-12-801522-3. DOI: `https://doi.org/10.1016/B978-0-12-801522-3.00006-9`. URL: `http://www.sciencedirect.com/science/article/pii/B9780128015223000069`.

[Zhu+17] Qiaoxi Zhu et al. "robust acoustic contrast control with reduced in-situ measurement by acoustic modeling". In: *journal of the audio engineering society* 65.6 (June 2017), pages 460–473. DOI: `https://doi.org/10.17743/jaes.2017.0016`.