



**HAL**  
open science

# Personalised low-frequency sound reproduction in a car cabin

Lucas Vindrola

► **To cite this version:**

Lucas Vindrola. Personalised low-frequency sound reproduction in a car cabin. Acoustics [physics.class-ph]. Le Mans Université, 2021. English. NNT : 2021LEMA1013 . tel-03414935

**HAL Id: tel-03414935**

**<https://theses.hal.science/tel-03414935>**

Submitted on 4 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THESE DE DOCTORAT DE

LE MANS UNIVERSITE

ECOLE DOCTORALE N° 602  
*Sciences pour l'Ingénieur*  
Spécialité : « *Acoustique* »

Par

**Lucas VINDROLA**

## **Personalised low-frequency sound reproduction in a car cabin**

Thèse présentée et soutenue à Le Mans, le 19 / 05 / 21

Unité de recherche : Laboratoire d'Acoustique de l'Université du Mans

Thèse N° : 2021LEMA1013

### **Rapporteurs avant soutenance :**

Philippe-Aubert Gauthier  
Jens Ahrens

Professeur, Université du Québec à Montréal.  
Associate Professor, Chalmers University of Technology

### **Composition du Jury :**

Président : Alexandre Garcia  
Examineurs : Veronique Larcher  
Philippe Herzog

Professeur, Conservatoire National des Arts et Métiers.  
Director AMBEO Immersive audio, Sennheiser  
Directeur général, Arteac - Lab

Dir. de thèse : Bruno Gazengel  
Co-dir. de thèse : Manuel Melon

Professeur, Le Mans Université  
Professeur, Le Mans Université

### **Invité**

Jean-Christophe Chamard

Ingénieur de recherche, Stellantis



# Résumé

L'étude de la génération de bulles sonores est un sujet de grand intérêt depuis plus de vingt ans, l'industrie automobile étant l'un des principaux domaines d'application. L'objectif de cette thèse est d'étudier la faisabilité d'un système permettant de générer deux bulles (appelées zones) dans un habitacle automobile, les deux zones ayant des contenus sonores différents n'interagissant pas entre eux. Ce système doit être capable de maintenir ses performances pour des fréquences inférieures à 1 kHz, quelque soient les changements qui pourraient survenir.

La première partie de cette thèse porte sur la conception d'un système capable d'obtenir un contraste acoustique (rapport d'énergies potentielles acoustiques entre les zones) d'environ 30 dB, tout en maintenant une faible erreur de reproduction dans la zone d'écoute. Un système à huit haut-parleurs a été proposé, consistant en deux appui-têtes comportant chacun quatre haut-parleurs. Le prototype a été testé expérimentalement dans trois environnements acoustiques différents : conditions anéchoïques, conditions anéchoïques avec panneaux réfléchissants et habitacle de voiture. Grâce à l'algorithme de « pressure matching », le prototype a permis d'obtenir un contraste acoustique moyen de 34,7 dB et une erreur de reproduction moyenne de  $1.7 \times 10^{-4}$  sur la largeur de bande étudiée dans l'environnement le plus complexe (habitacle automobile).

La deuxième partie de cette thèse aborde la variabilité acoustique inhérente à l'habitacle. À l'aide du prototype susmentionné, trois modifications ont été étudiées expérimentalement : la position des sièges, l'élévation de l'appui-tête et l'ajout de passagers sur les sièges arrières. Les résultats montrent que le prototype maintient ses performances quelle que soit la modification introduite si un ensemble de filtres spécifiques est utilisé pour chaque configuration.

Deux alternatives ont été étudiées pour obtenir ces filtres. L'une consiste à caractériser toutes les modifications possibles avant l'utilisation du véhicule (solution de type « base de données »), l'autre à mesurer la réponse de l'habitacle à l'aide de microphones pour recalculer les filtres en cours d'utilisation (solution de type « re-mesure »). Afin de rendre ces solutions plus adaptées à l'application automobile, la méthode « forced pressure matching » (FPM) a été développée. Cette méthode suppose que certains filtres sont imposés et permet de repenser la méthode du « pressure matching ». L'utilisation de la méthode FPM permet de réduire de 25% la quantité de stockage de la base de données et le temps de calcul d'environ 21%, ceci sans compromettre les performances.

Une troisième alternative propose d'utiliser un algorithme de filtrage adaptatif, la méthode « filtered-x least mean squares » (FxLMS) pour s'adapter aux modifications. La caractéristique essentielle de sa mise en œuvre est que, pour calculer les filtres, il suffit d'avoir une connaissance imparfaite de la réponse de l'habitacle de la voiture (modèle du système). Dans cette thèse, le modèle du système est estimé en mesurant la réponse de l'habitacle une seule fois en une position. Cette réponse permet d'obtenir les coefficients initiaux des filtres nécessaires à l'algorithme FxLMS (à l'aide l'algorithme de « pressure matching »). Les résultats obtenus montrent que la méthode de filtrage adaptatif est capable de maintenir les performances du système quelle que soit la modification acoustique introduite. Enfin, une étude approfondie de la réduction du nombre de microphones installés dans la voiture montre que le système reste performant avec seulement huit microphones d'erreur, soit deux par oreille.



# Acknowledgements

First and foremost, I would like to thank Manuel Melon, Bruno Gazengel and Jean-Christophe Chamard for their continuous support, knowledge and trust. I will forever cherish your patience, integrity and professionalism.

I would like to acknowledge the opportunities that the university of Le Mans gave me, from the IMDEA to working in the LAUM. It has been a privilege to be part of this team of talented researchers. Many thanks to Laurent Simon, Antonin Novak, Olivier Dazel, and Gilles Tissot for the invaluable discussions and brainstorming-sessions, many ideas from these encounters are reflected in this manuscript. Special thanks James Blondeau for the hours spent with me in the anechoic chamber installing and tuning the system.

I would like to thank my good friend Colas, for his help during the measurement sessions and for sharing many « *mates* » with me.

My heartfelt thanks to Manuel Melon and his family, for welcoming me in their house and making me feel part of the family.

I would also like to acknowledge the opportunity given to me by Stellantis (former PSA Groupe) to work on what I am passionate about. I would like to express my gratitude to Ladimir Prince for his interest and support of the project since the beginning.

I would like to express my deepest gratitude to the SCFH team, for welcoming me, supporting me and giving me a place to develop my research. Especially to the Hispanic-team, Luciano, Giacomo, Joaquim and Seb who made me feel at home from day one.

My sincere thanks goes to everyone at Velizy's workshop, in particular to Jean Lecuyer for his impeccable work, his amiability and for saving my life more than once.

I am also grateful to Laurent Simon and Alexandre Garcia for their review and feedback during the CSI sessions.

I would like to also thank all the members of the jury, in particular the reviewers Philippe-Aubert Gauthier and Jens Ahrens for taking the time to evaluate this manuscript.

*Quiero agradecer en especial:*

*A mis viejos y a mi hermano por su cariño y apoyo incondicional, esto es el fruto de sus esfuerzos.*

*A mis abuelos que ya no están, especialmente a vos Cuquita.*

*Y a la familia que voy adoptando en el camino, Rapo, Masa, Santi, Fio, Daniil, Fonti, Pepe, Lucas, Castaños y muchos más...*

Getting this far has been nothing but a team effort,  
it would not have been possible without you all...



# Contents

<b>List of Symbols and Abbreviations</b>	<b>IX</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Industrial requirements and constraints	1
1.1.1 Frequency band	2
1.1.2 Zones to control	2
1.1.3 Environment	2
1.1.4 Performance requirements	3
1.1.5 Feasibility constraints	3
1.2 Context	3
1.3 Summary and outline of the thesis	4
<b>2 Literature Review</b>	<b>5</b>
2.1 Problem definition	5
2.2 Methods	5
2.2.1 Directivity Control	6
2.2.2 Analytical sound field synthesis	12
2.2.3 Methods based on constrained optimisation	14
2.2.4 Conclusions	17
2.3 PSZ as a constrained optimisation problem	18
2.3.1 Deterministic approach: expressing the problem in the time and frequency domains	18
2.3.2 Metrics	20
2.3.3 The cost function	21
2.4 Perceptual evaluation	28
2.5 The procedure: synthesis of the method	29
2.6 Summary	31
<b>3 Invariant acoustics &amp; system definition</b>	<b>33</b>
3.1 The starting point: system definition	34
3.2 Free-field conditions	35
3.2.1 Complexity №1: Monopoles in free-field	36
3.2.2 Complexity №2: Prototype in anechoic chamber	43
3.3 Controlled reflections	47
3.3.1 Complexity №3: Monopoles with infinite-baffles	47
3.3.2 Complexity №4: Prototype with reflective panels	51
3.4 Car cabin	56
3.4.1 Complexity №5: Prototype in a cabin	57
3.5 Notes on the resulting filters: frequency or time formulation	64
3.5.1 Time domain	64
3.5.2 Time or Frequency?	67

3.6	Summary . . . . .	69
<b>4</b>	<b>Variable acoustics</b>	<b>71</b>
4.1	Literature review extension: PSZ under variable acoustics . . . . .	71
4.2	A library of impulse responses . . . . .	74
4.3	Sensitivity evaluation . . . . .	75
4.3.1	Seat position change . . . . .	76
4.3.2	Headrest position change . . . . .	79
4.3.3	Passenger configurations . . . . .	80
4.3.4	Summary . . . . .	81
4.4	Database solution . . . . .	81
4.4.1	Interpolation: saving time and storage . . . . .	82
4.4.2	Database conclusions . . . . .	83
4.5	Re-measurement solution . . . . .	83
4.5.1	Microphone-reduction . . . . .	84
4.5.2	Microphone-reduction extension: Virtual Microphones . . . . .	87
4.5.3	Re-measurement conclusions . . . . .	91
4.6	Forced-pressure matching . . . . .	94
4.6.1	Forced Pressure Matching theory . . . . .	94
4.6.2	The forced filters, the desired and their relationship . . . . .	96
4.6.3	The FPM performance: Complexity N°5 metrics . . . . .	97
4.6.4	The FPM performance: Extension to a variable car cabin . . . . .	100
4.6.5	Discussion FPM . . . . .	101
4.7	Summary . . . . .	102
<b>5</b>	<b>Adaptive approach</b>	<b>105</b>
5.1	Literature review extension: PSZ and the adaptive algorithms . . . . .	105
5.2	Adaptive PSZ: FxLMS theory . . . . .	108
5.3	FxLMS definitions . . . . .	110
5.3.1	FxLMS: step size parameter and convergence characteristics . . . . .	110
5.3.2	FxLMS: the initial coefficients and the plant model . . . . .	115
5.3.3	FxLMS: desired filter vector . . . . .	117
5.4	FxLMS variable acoustics: microphone reduction . . . . .	119
5.4.1	Forced FxLMS: ideas from previous chapters . . . . .	121
5.5	Different types of acoustic changes . . . . .	123
5.6	Discussion Adaptive PSZ . . . . .	125
<b>6</b>	<b>Conclusions and future work</b>	<b>129</b>
6.1	Conclusions . . . . .	129
6.1.1	Invariant acoustics . . . . .	129
6.1.2	Variable acoustics . . . . .	130
6.2	Future work . . . . .	132
	<b>Appendices</b>	<b>135</b>
	<b>A Headrests in cars</b>	<b>137</b>
	<b>B Prototype Schematics</b>	<b>139</b>
	<b>C Datasheet</b>	<b>141</b>
	<b>D Online results</b>	<b>145</b>

<b>E</b>	<b>Crossed Movements Car cabin</b>	<b>147</b>
<b>F</b>	<b>FxLMS: using simulated IRs</b>	<b>151</b>
	F.1 Flat responses . . . . .	151
	F.2 High-pass responses . . . . .	153
	<b>Bibliography</b>	<b>157</b>
	<b>Résumé étendu</b>	<b>165</b>



# List of Symbols and Abbreviations

ACC	Acoustic Contrast Control
ANC	Active Noise Control
<i>DZ</i>	Dark Zone
BACC	Broad-band Acoustic Contrast Control
BACCPM	Broad-band Acoustic Contrast Control with Pressure matching
<i>BZ</i>	Bright Zone
CO	Constrained Optimisation
CS	Combined Solution
FIR	Finite Impulse Response
FPM	Forced Pressure Matching
FRF	Frequency Response Function
FxLMS	Filtered-x least-mean-squares
FxRLS	Filtered-x recursive-least-squares
HOA	Higher order ambisonics
IR	Impulse Response
LMS	Least-mean-squares
LSM	Least-squares method
PSZ	Personal Sound Zones
PM	Pressure Matching
<i>Ref</i>	Reference Configuration
SDM	Spectral Division Method
SFS	Sound Field Synthesis
SPL	Sound Pressure Level
TF	Transfer Function
WFS	Wave Field synthesis
<i>AC</i>	Acoustic Contrast
<i>AC<sub>G</sub></i>	Acoustic Contrast Gain
CP	Calibration Position
Cond.	Condition Number
<i>D</i>	Driver
<i>E<sub>ff</sub></i>	Effort
<i>E<sub>rr</sub></i>	Normalised Reproduction Error
<i>E<sub>rrG</sub></i>	Reproduction Error Gain
EP	Evaluation Position
<i>h<sub>k,l</sub></i>	Impulse response between microphone <i>k</i> and source <i>l</i>
<i>P</i>	Passenger
$\epsilon$	Potential energy
<i>RAC</i>	Reference Acoustic Contrast
<i>SNR</i>	Signal to Noise Ratio

$TIR$	Target to Interferer Ratio
$\delta[n]$	Unit sampe function
$\mathcal{L}$	Lagrangian Equation
$c_0$	Speed of sound
$\mathbf{0}$	Vector of zeros
$\lambda$	Wavelength

# Chapter 1

## Introduction

Nowadays, cars are not only means of transportation but also items of luxury and comfort. Many branches of engineering are involved in the development of cars, not only to improve safety and reduce emissions, but also to enhance the travel experience.

Several sound producing technologies<sup>1</sup> have been incorporated in cars, both to assist the driver as well as to entertain the passengers. The problem is that not all these technologies are necessarily targeted to every passenger, which can lead to distraction and/or annoyance. Hence, the focus of this thesis is to design a sound system,<sup>2</sup> capable of reproducing dedicated audio programs to different passengers in a car.

The individualisation of audio programmes has been a topic of great interest since it was first studied by Druyvesteyn and Garas in 1997 [1]. This topic can be found both in the literature and in the industry with many different names: personal sound [1], acoustically bright and dark zones [2], personal audio [3, 4, 5, 6], multizone sound reproduction [7], individual sound zones [8], separated sound technology [9], private audio system [10], among others. Nevertheless, all these terms aim towards the same goal:

To allow different persons to listen to different programs in one environment, with minimum annoyance from the other programs, without using headphones or division panels.

Due to the fact that these systems are designed to generate different audio signals in spatially confined regions,<sup>3</sup> in this thesis the term personal sound zones (PSZ) is adopted.

Although the focus of this dissertation is the automotive industry, the generation of PSZ has a much wider field of application. For example, the need of having different levels between young members of a family and old members while watching TV to compensate hearing loss [11], to avoid the use of headsets in aircrafts [3], to create independent zones of music in public places underneath parasols to decrease ambient noise [12], in a domestic room [13], etc. Nevertheless, probably the most studied PSZ application is the car industry [14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24].

### 1.1 Industrial requirements and constraints

One of the first challenges of this thesis is to establish not only the desired results but also the design restrictions that determine the feasibility of the system.

The automotive application inherently defines a number of constraints and requirements that cannot be overlooked when designing the system. And yet, even among different publications involving cars, one can find differences in the system's requirements and constraints. These differences, can be either overly-restrictive or insufficiently-restrictive. Therefore, one proceeded to define the guidelines of this dissertation.

---

<sup>1</sup>For example: radio, driving assistance, hands-free phone, video players, GPS, etc.

<sup>2</sup>Including loudspeakers, signal processing units, amplifiers, microphones, etc.

<sup>3</sup>An area or a volume in a specific environment

### 1.1.1 Frequency band

The frequency band of interest of a car sound system in theory should cover the complete audible band (20 Hz - 20 kHz). However, the dimensional aspects of the sound system<sup>4</sup> and the wavelength of sound defines the working range of the different available technologies [1] (this is re-addressed in detail in Chapter 2).

In this dissertation, the focus is put in generating PSZ for the frequency band between 100 Hz to 1 kHz. Moreover, it is assumed that the higher part of the spectrum is controlled by a different technologies *e.g.* beamforming [22, 20], acoustic reflectors [25], etc.

### 1.1.2 Zones to control

The generation of PSZ implies to define the zones one wants to control.



Figure 1.1: Schematic: Car cabin with five zones.

The most typical zone-definition described in the literature is to consider both front seats ( $Z_1 + Z_2$ ) as one big zone and ( $Z_3 + Z_4 + Z_5$ ) as another big zone (see Fig. 1.1) [26, 19, 17, 16, 14]. Additionally, other studies have considered a diagonal front-back problem where  $Z_2$  is one zone and  $Z_3$  as the other zone [27, 23].

For simplicity, this thesis focuses on controlling two zones, the front-passenger ( $Z_1$ ) and the driver ( $Z_2$ ) without controlling the rear seats. Similar choice of control zones can be found in Ref. [3] (in an anechoic chamber and in a small room), or in Refs. [22, 20] (in car cabins) to control a higher frequency band. Although the extension of a multi-zone system (*e.g.* Refs. [24, 28, 21]) is out of the scope of this thesis, the conclusions and methodologies reviewed in this dissertation can be extrapolated to these larger systems.

### 1.1.3 Environment

The car cabin can be seen as a complex acoustic environment. It is made of highly reflective surfaces (glasses), highly absorbent surfaces (seats), diffractive elements, etc. Furthermore, each model has different optionals/configurations (different fabrics, different materials on the surfaces, glass-ceilings, etc.) that can affect the response of a sound system.

Moreover, the car environment is a user-modifiable environment which further-increases the complexity. For example, the relative location of the zones can be chosen by the user (the seat and the headrest are arranged in a different position depending on the height of the driver), the number of passengers in the car

<sup>4</sup>Size of the system, separation between sources, separation between the listening zones, size of transducers, etc.

varies, etc. Each modification in the environment can result in a modification of the acoustic environment, which might change/affect the system’s performance.

Therefore, one of the key requirements of this dissertation is to be able to perform under all these different acoustic conditions, even if some of these might not be possible to accurately estimate in advance.

#### 1.1.4 Performance requirements

The PSZ experience has the unique feature of being an audio-on-audio interference rather than a wide-band noise-on-audio interference. In other words, an “interference” audio programme interferes upon a “target” audio programme [29]. Therefore, one of the main requirements is to reduce the interference audio (coming from the other zone) as much as possible.

Moreover, the overall quality of the PSZ listening experience can be considered as the result of two effects, the presence of the interferer program and the degradations of the target audio [30]. Additionally, with the advent of autonomous vehicles, the entertainment system is becoming a key feature in the car’s cockpit with different sound spatialisation technologies, *e.g.* immersive-sound [31], spatial rendering technologies [32], etc.

Consequently, another important requirement is to maintain the same reproduction characteristics as if the PSZ system was not implemented (or as close as possible). In other words, the target programme is the system’s natural response (this is re-addressed in detail in Chapter 3).

To sum up, it is necessary to reduce the interference programme as much as possible (to avoid the audio-on audio interference) without altering the reproduction capabilities of the system. These two requirements are loosely referred to in this thesis as the “acoustic separation” and the “transparency” respectively.

#### 1.1.5 Feasibility constraints

Lastly, inherent to the car application, there are other constraints that need to be taken into consideration when designing the PSZ system.

On the one hand, the cost and weight added due to the use of numerous independent channels is a design constraint that cannot be ignored. Therefore it has been imposed as a limit the use of less than ten independent audio channels. Despite its obviousness, it is important to highlight that the size of the system needs to be realistic for the car application, *i.e.* it should be possible to implement the solution in a car cabin.

On the other hand, another important parameter that needs to be held in check is the extra electrical power required by the generation of PSZ.

## 1.2 Context

This thesis is part of the industrial agreements for training through research (CIFRE), grant number 2017/1495. The CIFRE convention was funded by the French ministry of higher education, research and innovation and implemented by the national association for research and technology (ANRT).<sup>5</sup>

Moreover, the work presented in this thesis is part of a larger innovation project in collaboration between Stellantis (former PSA Groupe) and the acoustic laboratory of Le Mans university (LAUM).

Stellantis is a leading global mobility player guided by a clear mission: to provide freedom of movement for all through distinctive, appealing, affordable and sustainable mobility solutions. The group offers a full spectrum of choice from luxury, premium and mainstream passenger vehicles to pick-up trucks, SUVs and light commercial vehicles, as well as dedicated mobility, financial, and parts and service brands.<sup>6</sup>

The LAUM, is a joint research unit of the University of Le Mans and the CNRS (UMR 6613). The laboratory employs about 160 people (teacher-researchers, researchers, technical and administrative staff,

<sup>5</sup>[www.anrt.asso.fr](http://www.anrt.asso.fr)

<sup>6</sup>[www.stellantis.com](http://www.stellantis.com)

doctoral students, post-doctoral students and guests). The laboratory's activities are mainly focused on the acoustics of the audible, but in recent years the laboratory has integrated new research themes in the field of vibration and ultrasound.<sup>7</sup>

### 1.3 Summary and outline of the thesis

Lastly, the requirements/constraints governing this research are summarised in the following list:

- 1) Bandwidth of interest: below 1 kHz.
- 2) Zones: two front seats ( $\approx 73$  cm apart).
- 3) "Sufficient" attenuation of the interference audio programme.
- 4) Transparent system.
- 5) Variable car cabin.
- 6) Number of independent channels: less than 10.
- 7) "Acceptable" power consumption.

The list above establishes the framework used throughout this thesis in order to design and evaluate the performance of the PSZ sound system.

#### Outline of the thesis

Firstly, chapter 2 reviews different technologies that can be used for the generation of PSZ. Each technology is put in context with the application in hand (requirements and constraints). In addition, the metrics to evaluate the system are chosen and the procedure to generate PSZ are reviewed.

In chapter 3, one goes through the design and implementation of the system in different acoustic environments.

The prototype designed in the previous chapter is tested in chapter 4, when different modifications in the car-cabin are introduced. The goal of these experiments is twofold: firstly, to see if the system is capable of achieving the wanted performance under all these modifications, secondly, to study its implementation in a variable car cabin.

In Chapter 5, the possibility of implementing adaptive algorithms for PSZ is introduced to maintain the system's performance under changes in the acoustic conditions.

Lastly, in Chapter 6, the main conclusions are summarised, compared and the future perspectives are introduced.

---

<sup>7</sup>laum.univ-lemans.fr

# Chapter 2

## Literature Review

### 2.1 Problem definition

The simplest scenario to study the generation of PSZ is to control the sound pressure over two spatially confined zones  $Z_1$  and  $Z_2$ . The aim is to reproduce an audio program  $x_1[n]$  over the zone  $Z_1$  (referred to as the bright zone  $BZ$ ) whilst minimizing it in a different zone  $Z_2$  (referred to as the dark zone  $DZ$ ) (see Fig. 2.1).

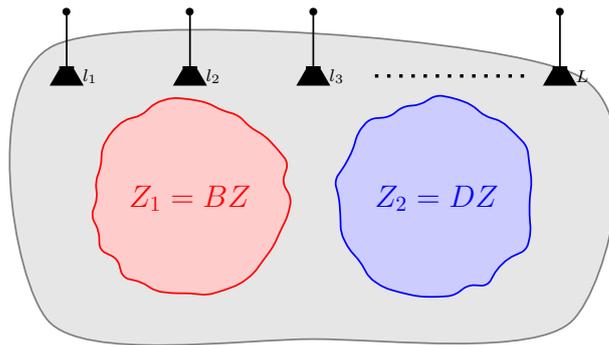


Figure 2.1: PSZ simplest case general schematic, one bright zone and one dark zone,  $L$  acoustic sources.

Therefore, using  $L$  acoustic sources -*e.g.* loudspeakers, vibrating panels, piezoelectric transducers, etc.- the aim of the PSZ is to control the interferences between these sources in order to generate the aforementioned  $BZ$  and  $DZ$ . As it is usually done in the literature, when more than one PSZ is wanted, the system responses can be simply superposed; *e.g.* if a different audio program  $x_2[n]$  is being played in  $Z_2$  whilst is being minimised in  $Z_1$ .

### 2.2 Methods

When one refers to PSZ, one directly associates to the methods based on constrained optimisation. Also referred to as active control of sound [1], active reproduction of sound [14], mutizone sound control [33], etc.

Although this report's main field of study deals with these methods, there are several other technologies that can be used towards the same goal. Due to the large differences between the applications of PSZ -like mobile devices [34], television arrays [11], aircraft [3], automotive [16], rooms [13], etc.- these methods are worth reviewing and passing them through the sieve of this application.

Due to the fact that not all these methods are necessarily applied in the same conditions (size/distribution of zones, environment, number of available sources, etc.), it is not so simple to compare among different

published works. Therefore, the focus of this analysis is placed on the theoretical assumptions and/or limitations of the different methods and if these assumptions are too restrictive for the studied application.

In the following, one can find different methods/technologies that can be used (in some cases have been explicitly used) for the generation of PSZ. These are divided in three different categories according to the criteria used for their design: directivity control, synthesis of a pressure field or optimisation of a cost function.

### 2.2.1 Directivity Control

One possibility for the generation of PSZ is to control the directivity of the sources. In other words, to maximise the direct sound to the *BZ* and minimise the direct and reverberant field to the *DZ* [33].

A few definitions need to be introduced before continuing with the analysis.

- 1) Directivity: The directivity of a sound source refers to the manner in which the measured or predicted sound pressure, at a fixed distance  $r_0$  from the source, varies with angular position  $\alpha$  [35] at a specific frequency, in a specified plane -usually normalised on-axis ( $\alpha = 0$ ) [36].

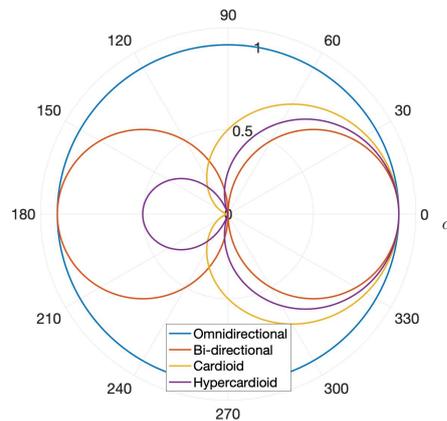


Figure 2.2: Theoretical directivity patterns: Omnidirectional (blue), Bi-directional (red) Cardioid (yellow) and Hypercardioid (purple).

For example in Fig. 2.2 one can observe four examples of theoretical directivity patterns that can be found in the literature.

- 2) Lobes (From Balanis 2012 [37]): various parts of a radiation pattern can be referred to as lobes. A radiation lobe is a portion of the radiation pattern bounded by regions of relatively weak radiation intensity. There are major lobes, minor lobes, side lobes and back lobes. To maintain the analysis as brief as possible one is going to introduce only two.

The major lobe, also referred to as main beam, is defined as the lobe containing the direction of maximum radiation. A side lobe, is a radiation lobe in any direction other than the main lobe.

- 3) Near-field vs. Far-field: the far field is going to be defined as the spatial region where the direct-path pressure amplitude varies inversely with distance from the source. Whilst the near-field, is the spatial region complementary to the far-field [38].

These definitions will help to analyse and compare the ideas behind the different methods.

### 2.2.1.1 Radiating elements

The first “family” of the directivity control methods is based on modifying the geometrical characteristics, the material properties or the excitation of the radiating elements in order to generate a directive source.

#### 2.2.1.1.1 Wave-type sources

The directivity pattern of a loudspeaker in a closed box, at low frequencies where the wavelength ( $\lambda$ ) is much larger than the radiating elements, can be assumed to be almost omnidirectional (see Fig. 2.2).

However, as  $\lambda$  becomes comparable to the dimensions of the radiating surfaces, the source becomes more directive [39]. Therefore, it is possible to control the directivity of a source by designing the radiator<sup>1</sup> in such a way that its size is comparable to the wavelength [40]. In other words, design the radiator having in mind that its directivity depends on the sound emanating from the elements of its radiating surfaces [40].

Many examples can be found in the industry of wave-type sources, for example the acoustic reflector patented by Bang & Olufsen in 2004 [25] (see left in Fig. 2.3), the parabolic reflector [41] (see right in Fig. 2.3), flat panel speakers [41], etc.

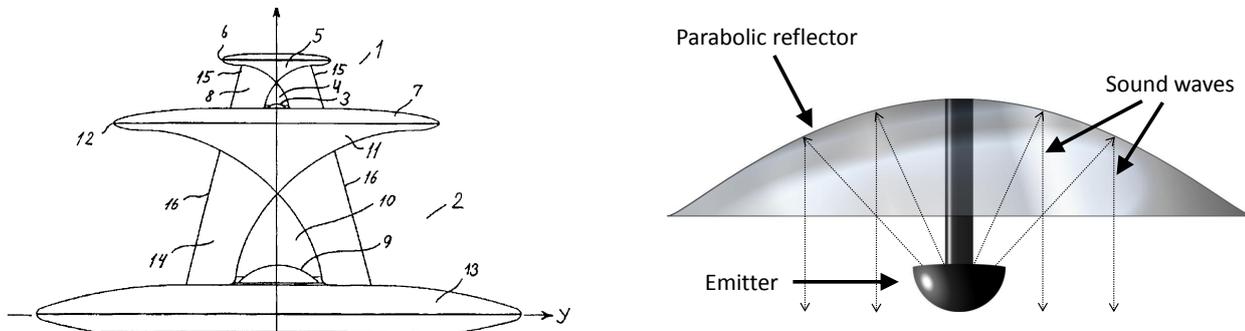


Figure 2.3: Examples of wave type sources: Acoustic reflector schematic (left) [25], parabolic reflector schematic (right) [41].

However, for the system to become directive, its dimensions need to be comparable to  $\lambda$ , which in the bandwidth of interest of this thesis, results in very large systems ( $\lambda \simeq 3.4$  m at 100 Hz) making the wave-type sources not viable for the studied frequencies. Nevertheless, it is worth mentioning that, these type of sources are interesting candidates to control higher frequencies.

#### 2.2.1.1.2 Metamaterials

Metamaterials are rationally designed composites built of mesoscopic<sup>2</sup> resonant inclusions placed in a host medium [42]. These inclusions, being smaller than the wavelength exhibit collectively an on-demand property which allows novel wave manipulation functionalities going beyond the conventional wisdom of refraction laws [42]. Changing the size and geometry of these resonators one can modify the acoustic properties of the metamaterial and its operation frequency [43].

Among the applications of metamaterials is the control the directivity of a sound source by designing an inhomogenous impedance boundary, introducing sub-wavelength Helmholtz resonators [44, 45] or “ultra-high” resolution sound focusing [42].

As reviewed in Ref. [42], the “bottleneck” of this particular technology is to be able to design a solution applicable for a broad-bandwidth (as the one reviewed in this study) as well as the reduction of the system losses. Although some active metamaterials have been studied to overcome these limitations, their use is ignored in this study. For an extensive review of the state of the art of metamaterials see Ref. [42].

<sup>1</sup>The radiator can be the surface of the diaphragm, the mouth of a horn, size of the reflector, etc. [40]

<sup>2</sup>Intermediate size between microscopic and macroscopic

### 2.2.1.1.3 Actuators

A topic that has gained some interest in the car industry is to replace the conventional electrodynamic loudspeakers by exciting select surfaces of the vehicle interior using actuators [46, 47].

Actuators have been used to deliver low-frequency vibrations in the seats to generate private low-frequency audio [47]. In addition, one could extend this premise to find the best placing of the actuators in order to achieve a desired directivity [41]. For example in 2018 Jeon *et al.* [48] presented how to improve the performance of a multi-actuator panel with low-damping loss factor by optimising the exciter placing according to the modal coupling between the different actuators, work that was later expanded in Ref. [49]. Nevertheless, the implementation of this technology is ignored from this study.

### 2.2.1.2 Interferences between simple sources

Another approach to tailor the directivity of sources is to control the acoustic interferences between more than one source. It was stated by Beranek that the basic principles governing the directivity patterns of loudspeakers can be studied as combinations of simple sources [36], but also one can take this approach inversely, to control the directivity pattern controlling a combination of simple sources.

Therefore, the radiation of the acoustic sources<sup>3</sup> is approximated as the one of simple sources (usually a monopoles in free-field) and their directivity depends upon the difference, in amplitude and phase, in the generated sound pressure between these simple sources [40].

#### 2.2.1.2.1 Phase-shift sources

One example of this principle is to use the radiation of the rear-side of a dynamic loudspeaker diaphragm as a second source, an idea that according to Holmes dates back at least to 1950 [50].

In other words, one can consider both “sides” of one loudspeaker as two monopoles separated by a distance  $\Delta_l$  with volume velocities  $Q_1$  and  $Q_2$  for the front and rear sides respectively (see Fig. 2.4a). As the pressure at an observation point ‘o’ is the vectorial addition [36], it is possible to control the directivity by modifying the relative phase between the volume velocities  $Q_1$  and  $Q_2$ .

Therefore, placing the loudspeaker in an enclosure with a rear opening where a resistive material of value  $R_A$  is placed as in Fig. 2.4b it is possible to modify the relative volume velocity between the two monopoles [50].

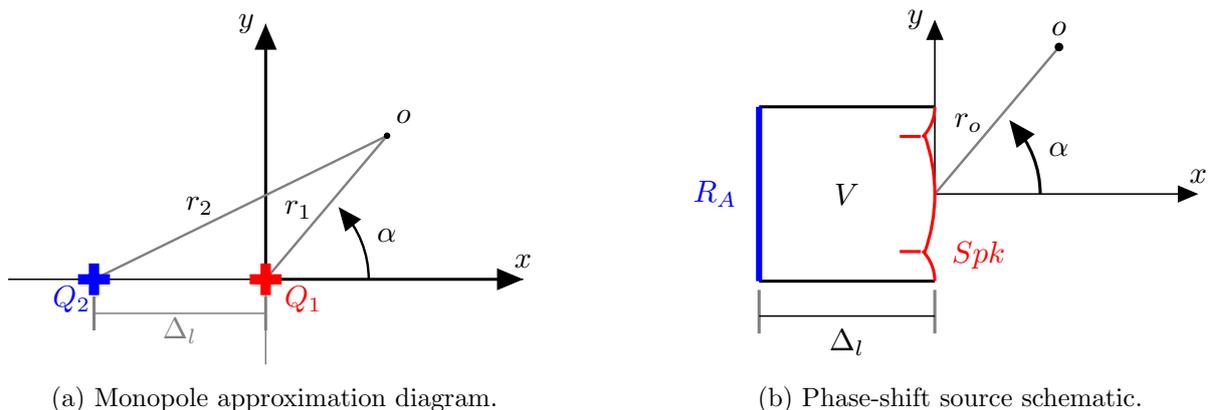


Figure 2.4: Phase-shift source diagrams.

Depending on the values of  $R_A$ , the size of volume  $V$  and the separation between the sources  $\Delta_l$  it is possible to obtain different radiation patterns. For example, to obtain a bidirectional radiation pattern

<sup>3</sup>It is important to clarify that in this chapter the number of acoustic sources not necessarily refers to the number of transducers.

$R_A = 0$  and to obtain a cardioid-like directivity  $R_A = \rho c_0 \Delta_l V^{-1}$ . This approach is also known as the “acoustic resistance-box” [50]

Moreover, one of the challenges of the methods discussed above is the generation of two or more “independent” sources from a single transducer. However, it was already mentioned that a loudspeaker in a closed-box behaves at low-frequencies as a monopole source, therefore, by placing two sources close to each other it is possible to calculate the relative phase-shift in order to obtain the wanted directivity pattern [39, 40]. This procedure is well-known in public address in order to avoid increased reverberation, reflections or to only radiate towards the audience and the scenario [39].

To sum up, it is possible to re-interpret the single-transducer phase-shift with independent closed-box loudspeakers and calculate the best combination of driving signals as a function of their separation ( $\Delta_l$ ) and the wanted directivity.

Lastly, there are two main assumptions that need to be fulfilled in order to achieve directive patterns (cardioid/hypercardioid in Fig. 2.2): the long-wave situation (Low-frequency assumption) and the far-field assumption [50, 39].

On the one hand, the long-wave situation assumes that  $\lambda/2\pi \gg \Delta_l$ , which, due to the frequency band of interest of this thesis, is not so-restrictive and should be possible to fulfil. On the other hand, the far-field assumption demands that the observation distance  $r_o \gg \Delta_l$ , if not met, it results in a deterioration of the directivity-performance. As the aim of this report is to generate PSZ between the two front seats, this assumption is more restrictive for the application in hand. Nevertheless, it is necessary to mention that the use of phase-shift sources for PSZ in a car application was previously studied in Refs. [14, 15, 16]. However, in these previous studies, the zones were defined as the front and back seats, making the separation between them considerably larger than the studied case.

### Proximity-effect

Another example of using the phase-shift approach but in a slightly different way is the recent product patented by Premium Sound Solutions [51, 52] the “sound cocoon”. In this patent, the inventor proposes to install a dipole loudspeaker in the headrest (see Fig. 2.5), *i.e.* a loudspeaker with an open-back enclosure, or the passive phase-shift source when  $R_A = 0$ .

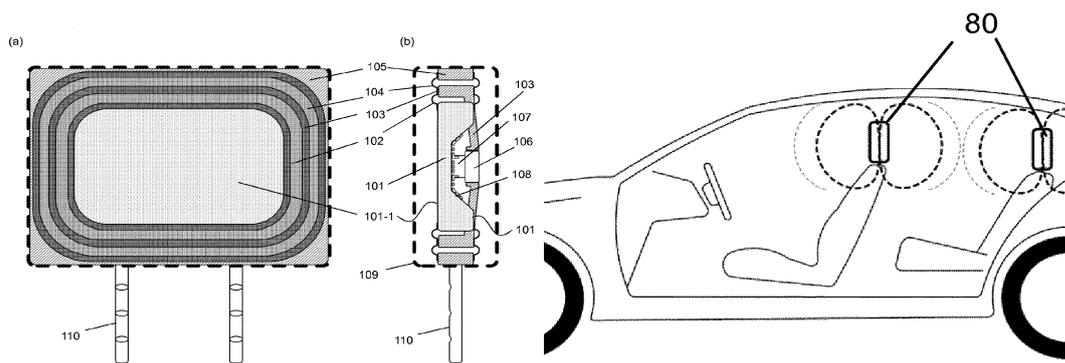


Fig. 8

Figure 2.5: Sound Cocoon: headrest schematic (left) and example of implementation in a car (right) from Ref. [51].

In the mid-high frequency (above 150 Hz for the application in Refs. [51, 52]), the open-back loudspeaker presents a bidirectional directivity (see Fig. 2.2). However, for “very-low” frequencies and away of the headrest, the SPL of the dipole drops considerably faster than a monopole whilst at a small distance of the headrest the SPL is almost unaltered (also known as the proximity effect) [51].

Although the sound cocoon can be interpreted as a phase-shift source when  $R_A = 0$  (in Fig. 2.4b), its principle of operation is based on the proximity effect and not in the resulting directivity pattern of the source in the far-field.

One of the key parameters of the sound cocoon is what the inventor refers to as “path length” or (the  $\Delta_l$  in Fig.2.4b), *i.e.* the equivalent distance between two out of phase monopoles. As  $\Delta_l$  increases, the frequency of application of this invention decreases. The theoretical analysis behind this idea can be found in Ref. [53].

Due to the fact that  $\Delta_l$  is limited by they size of the transducer, this type of solution is targeted to control the lowest part of the audible spectrum, *i.e.* frequencies no higher than  $\approx 150$  Hz [51].

### 2.2.1.2.2 Delay and sum beamforming

The main principle of controlling the interferences between simple sources can be extended from two sources (phase-shift sources) to  $L$  sources (see Fig. 2.6).

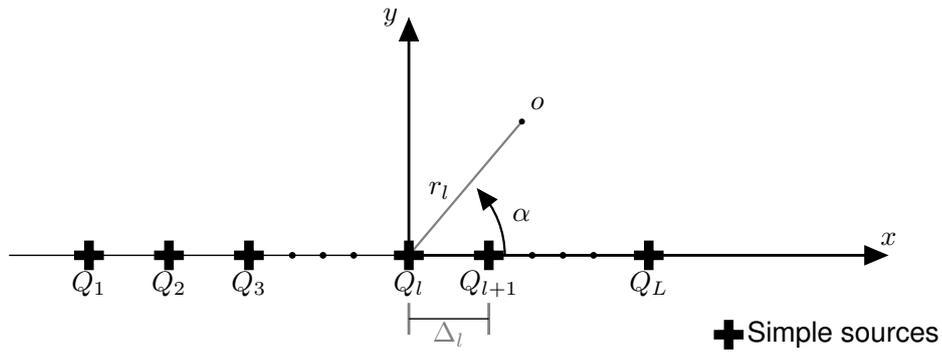


Figure 2.6: Beamforming schematic.

As the number of sources becomes larger, the directivity pattern becomes more complicated, therefore when more than two acoustic sources are present their directivity is studied in terms of beams or lobes, hence the beamforming distinction. Therefore, when one refers to beamforming one refers to the generation of a main-lobe or main-beam in the wanted direction while the generated side-lobes are “sufficiently” attenuated.

Therefore, by controlling the relative amplitudes and phases, between all the monopole’s volume velocities ( $Q_1, Q_2 \dots Q_L$ ) as a function of their spacing  $\Delta_l$  (equidistantly spaced or not [54]) and the wanted direction of the main lobe (defined by an  $\alpha_{main}$  towards the  $BZ$ ), it is possible to steer the array’s directivity in such a way that a main beam is generated in the  $\alpha_{main}$  direction and the nulls -and/or the attenuated side lobes- of the pattern point towards the  $DZ$  [55, 56].

This principle when employed using closed-box loudspeakers next to each other, is known as delay and sum beamforming, coherent summation array, etc. or if  $\alpha = 0$  (see Fig. 2.6) it is referred to as the end-fire array [56].

The multi-source interferences has also its single-transducer example, the so-called shotgun loudspeaker. This approach is based on one transducer radiating inside a pipe to which a series of holes have been drilled with a separation of  $\Delta_{holes}$  along its axis closed by an anechoic termination (see Fig. 2.7). As the sound wave passes along the pipe, each hole acts as an individual sound sources.

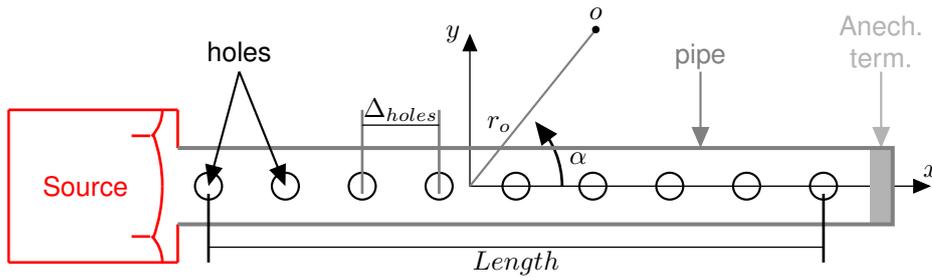


Figure 2.7: Shotgun loudspeaker schematic.

Because the output from each hole is delayed between each other -due to the propagation of sound inside the pipe-, if the delay between to holes is equal to  $\frac{\Delta_{holes}}{c_0}$ , the resultant array generates a main beam in the direction of the propagating wave (along the x-axis in Fig. 2.7 or  $\alpha_{main} = 0$ ) -this principle is the same as the one used in shotgun microphones [57]. Therefore, by properly adjusting the dimensions of the tube, the distance between holes, etc. it is possible to obtain the needed delay. The same approach can be employed using multiple ducts or hoses so that the sound from a single source can be conducted to several points along an array [56].

The use of delay and sum beamforming has been widely studied in applications such as RADAR, SONAR, Biomedical, etc. [58] and there is a vast literature of different improvements and modifications, multiple-beam steered arrays [59], array outlines to improve the frequency-dependence of the beam [54], etc.

Although the delay and sum beamforming is a sort-of generalisation of the phase-shift sources in Olson 1973 [40], it is defined as another type of wave-type source probably because of its assumptions and limitations. Firstly, the lowest operating frequency is determined by the overall size of the array;  $\lambda < \Delta_l(L - 1)$  needs to be met to achieve a directive response [60]. Secondly, to avoid spatial aliasing it is necessary to have at least two sources per-wavelength ( $\Delta_l \leq \frac{c_0}{2f_{max}}$ ) [54]. Thirdly, in order to achieve the wanted directivity-control it is necessary to be in the far-field of the source [57].

In this thesis, the high frequency limitation might not generate strong complications, as long as the number of sources is maintained small, the spacing between sources at 1 kHz should be feasible. On the other hand, due to the limited size of a car cabin it is not possible to obtain an array large-enough to achieve a directive response in the lowest part of the frequency band of interest.

Furthermore, another complication of these methods is the implementation in a confined environment such as the car cabin. In addition, the derivation of the phase-shift between the sources is estimated from the assumption that the closed-box loudspeakers are monopoles radiating in free-field. However, as the Schroeder frequency in a car compartment is typically around 300 – 440 Hz [61, 62], the modal response of the cabin should be taken into consideration.

### 2.2.1.3 Ultrasonic demodulation

The parametric loudspeaker or antenna is an alternative to generate directive sources. Its fundamental theory is based on the work of Westervelt in 1963 [63]. An ultrasonic signal (the carrier) is amplitude modulated with an audio signal and radiated from an array of ultrasonic transducers as finite amplitude waves. When the amplitude modulated ultrasound propagates, due to the non-linear characteristics of the medium, the audio signal is demodulated in the air. The main interest of these particular sources is that an extremely sharp directivity is achieved for the audio signal [64].

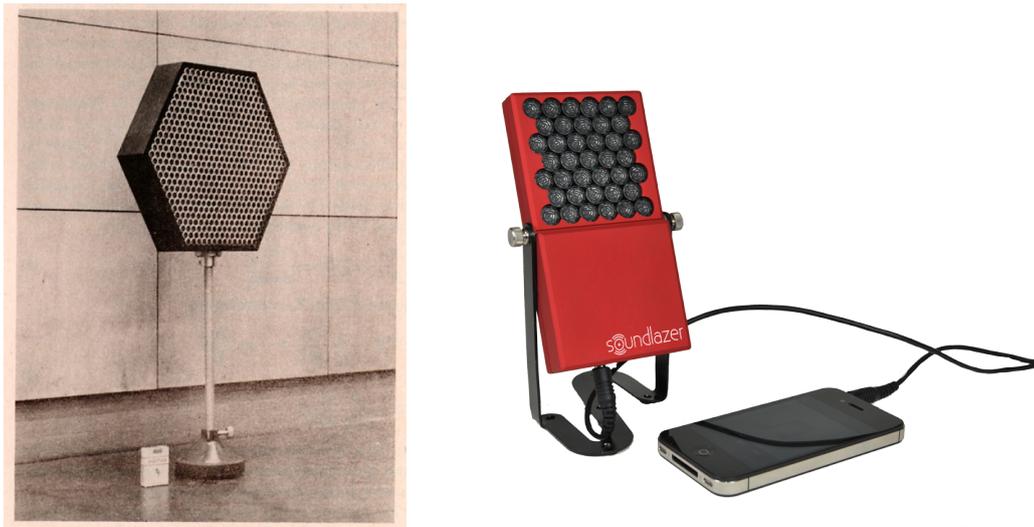


Figure 2.8: Parametric antenna: Yoneyama 1983 (left) [64] and Soundlazer (right) [65].

The parametric loudspeaker is usually implemented with several ultrasonic transducers, therefore many of the beamforming principles studied above could be applied in order to steer and/or generate multiple main beams [66]. However, the core of operation is still the ultrasonic demodulation.

Unfortunately, in order to achieve reasonable levels of the audio signal, the carrier sound pressure needs to be in extremely high SPL levels ( $\simeq 130$  dB SPL) which may cause unpleasant sensation, headache, fatigue and nausea. It has even been reported that long hours of exposure may cause threshold shift in hearing-loss [67]. These concerns, and the prolonged hours of exposure one might have in a car make the parametric solution to be unsuitable for this application.

## 2.2.2 Analytical sound field synthesis

One particular approach that needs to be reviewed in detail is the idea of sound field synthesis (SFS). The SFS problem can be posed as: *A given ensemble of elementary sound sources shall be driven such that the superposition of the sound fields emitted by the individual elementary sound sources makes up a sound field with a given desired properties over an extended area.* [68].

In other words, the aim of SFS is to recreate a desired pressure field in a specific area or volume surrounded by loudspeakers [69], also referred to as secondary sources (see Fig 2.9). The distinctive characteristic of these methods, from the the ones reviewed above, is the fact that a desired pressure field<sup>4</sup> is defined. This desired pressure field can be defined as the result of a virtual source located behind the array (as in Fig. 2.9), a source in front, but also by any other target field, for example, a plane wave field, or a multizone reproduction field [70].

In this section, one is focusing on the analytical SFS, a method based upon analytic formulations of the underlying spatially continuous physical problem, the pressure matching approach -usually listed among the SFS- is reviewed below.

<sup>4</sup>In this thesis, desired pressure field and target pressure field are used as synonyms.

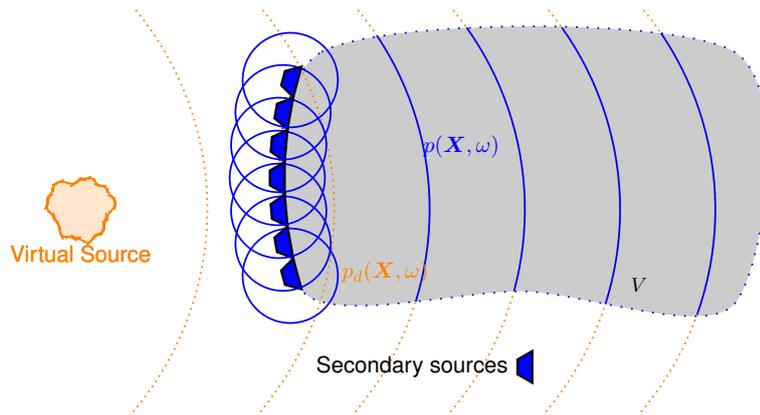


Figure 2.9: Sound field synthesis schematic: desired pressure field  $p_d(\mathbf{X}, \omega)$  and synthesised pressure field  $p(\mathbf{X}, \omega)$ , being  $\mathbf{X}$  the spatial coordinate vector and  $\omega$  the angular frequency.

Probably the two best-known methods concerning analytical SFS are wave field synthesis (WFS) and higher order ambisonics HOA. Although it is possible to use a general formulation to compare both methods [71], one is going to use their original formulations.

The basis of WFS is formed by Huygens's principle, that a wave front may be considered as a secondary source distribution [72, 73]. Therefore, expressing the synthesized pressure field using Kirchoff-Helmholtz integral equation, which states that a continuous distribution of appropriately driven secondary monopole and dipole sources arranged on the boundary of the desired listening area is capable to reproduce any virtual wave field inside that area [72, 74].

The derivation of the driving signals assumes a continuous distribution of sources in an anechoic reproduction environment, which for practical reasons needs to be adjusted [72, 73]. Moreover, the continuous distribution of sources is replaced by an arrangement of closed-box loudspeakers (monopole-like directivity) at discrete positions which means spatial sampling and truncation [72, 74, 75, 73], additionally, the anechoic environment is usually a common room with modes and reverberation [76, 72, 77].

On the other hand, the principle of HOA is based on matching the desired soundfield with the soundfields resulted from the secondary sources (i.e. loudspeakers) in terms of the harmonics-based expansion for 3-D (spherical harmonics representation) or 2-D (cylindrical harmonics representation) cases [74, 69]. It is important to mention that it is also possible to implement the same principle for planar and linear secondary source distributions, also known as spectral division method (SDM) [78]. Originally, the Ambisonics system reproduced the zeroth and first order spherical/cylindrical harmonics at a single point and was later extended to higher order as higher order Ambisonics (HOA) which enlarged the reproduction area [74, 69].

Although the analytical methods reviewed above are usually implemented in loudspeaker setups which -partially or fully- enclose the listening area, their principles can be extended to outward radiating setups [68], for example in beamforming-like applications [79, 80] or for the generation of evanescent-waves sources [81, 82].

For a more detail review and/or comparison between SFS methods refer to [68, 69, 71, 73, 74, 78].

The SFS realisation, requires a careful consideration of different practical aspects [78]. Most of the SFS techniques assume anechoic room, a requirement not easy to achieve in a car cabin. Although some researchers proposed different methods to compensate the influence of the room both actively and passively [75, 77, 83, 84, 85], the controllability of the reproduced sound might be deteriorated due to the cabin's acoustical modes [78]. Moreover, the quality of the sound reproduction system depends heavily on the quality of the individual loudspeakers used [78], quality that might not always be possible to achieve in a car, due to size of the transducer, its enclosure, weight, costs and also product-variability. In addition, another aspect that might restrict the use of the analytical SFS methods is the necessary processing capabilities to properly drive the sources (especially for HOA) [78].

Lastly, and probably the main complication for the studied application is that, in order to achieve an

accurate reproduction over an extended area, the analytical SFS methods usually require a large number of sources [78]. To sum up, it was already observed in Refs. [21, 86], that the major challenge of these analytical methods is to meet the restrictive assumptions -about the number of sources, their properties, their distribution, the size and shape of the zones, the characteristics of the environment, etc.-.

Although, the analytical SFS methods are not considered for the studied application, it is necessary to mention that many other authors have successfully applied these technologies in other PSZ applications [87, 88, 70, 74, 89, 90, 80, 91, 92] as well as in a concept-car using 62 loudspeakers [93, 14].

### 2.2.3 Methods based on constrained optimisation

All the methods reviewed up to now assume a mathematical knowledge of the physical problem as a function of the source position, their directivity, their propagation characteristics, the location of the zones, the environment, etc. Therefore, in order to derive the driving signals (or filters) a series of assumptions -*e.g.* anechoic environment, monopole sources, etc.- are made but are valid to a certain extent. Due to the fact that a car cabin is a complex acoustic environment (irregular shapes, very reflective and very absorbent surfaces distributed along the cabin), the channel restrictions, the frequency band of interest (modal response of the cabin, the wavelength  $3.4 \text{ m} \leq \lambda \leq 0.34 \text{ m}$ ) and also the variability between different car models (different car-sizes, different configurations, etc.) many of these assumptions are difficult to meet in the studied application.

Therefore, the last “family” of methods reviewed, and the core of this dissertation, is the constrained optimisation (CO).<sup>5</sup> These methods are based on sampling/discretising the zones to control with  $K$  microphones ( $k_1, k_2, \dots, K$ ) and representing the acoustical problem as the impulse responses ( $h_{k,l}$ ) of length  $I$  measured between the sources ( $l_1, l_2, \dots, L$ ) and the  $K$  microphones (see Fig. 2.10a). It is important to highlight that these methods leave a zone completely uncontrolled (grey area in Fig. 2.10a).

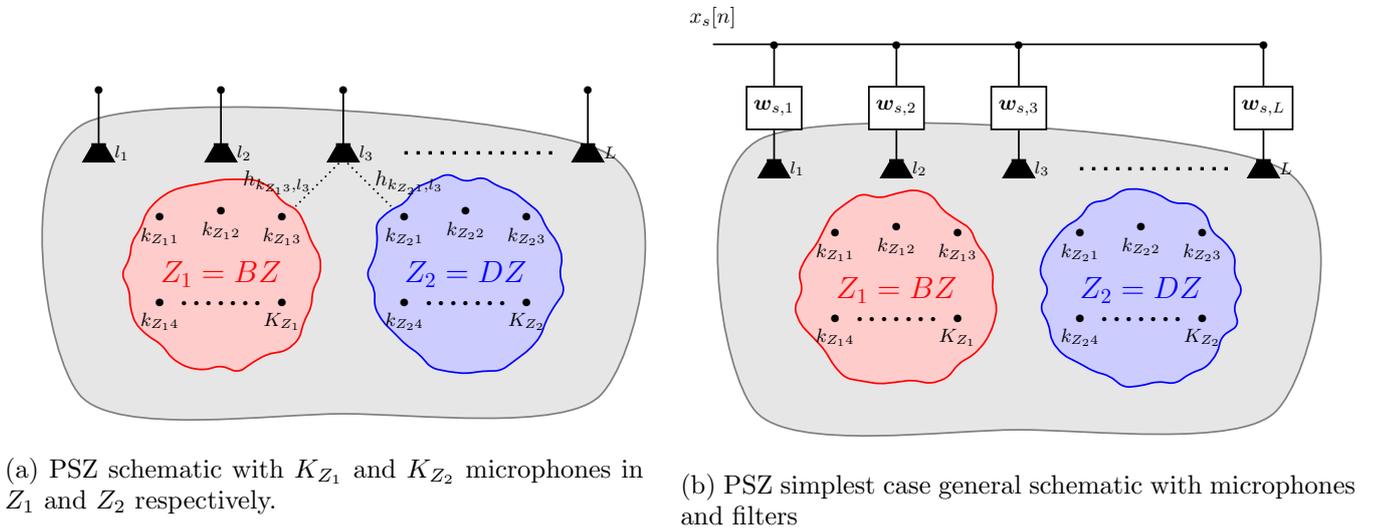


Figure 2.10: Constrained optimisation schematics, zones discretisation.

Therefore, if one assumes a perfectly linear system, the sampled pressure at microphone  $k$  ( $p_k[n]$ ) can be expressed for a signal  $s$

$$p_{k,s}[n] = \sum_{l=1}^L \sum_{i=0}^{I-1} h_{k,l}[i] u_{s,l}[n-i], \quad (2.1)$$

<sup>5</sup>The terms multipoint methods is also incorporated as a synonym, in spite of being used mainly to describe the CO associated to SFS [78]

being  $u_{s,l}[n]$  the signal driving the  $l$  source as in

$$u_{s,l}[n] = \sum_{m=0}^{M-1} w_{s,l}[m] x_s[n-m], \quad (2.2)$$

where  $M$  is the number of coefficients of the filters  $w_{s,l}$ . Therefore, replacing Eq. 2.2 in Eq. 2.1 yields

$$p_{k,s}[n] = \sum_{l=1}^L \sum_{i=0}^{I-1} h_{k,l}[i] \sum_{m=0}^{M-1} w_{s,l}[m] x_s[n-m-i]. \quad (2.3)$$

The CO methods calculate the “best” filter combination ( $\mathbf{w}_s = [\mathbf{w}_{s,1}^T, \mathbf{w}_{s,2}^T, \dots, \mathbf{w}_{s,L}^T]$ ) to be applied, usually as a feed-forward control technique [3, 94]. In other words, it obtains the filtered versions of  $x_s[n]$  ( $u_{s,l}[n]$  in Eq.2.2) to drive the  $L$  sources in such a way that the resulting pressure in the zones (Eq. 2.3) -through constructive and destructive interference [14]- results in the so-called *BZ* and *DZ* (see Fig. 2.10b).

The calculation of the optimal filters is based on numerical optimisation; expressing a cost function using the IRs and minimise/maximise it as a function of  $\mathbf{w}_s$ . This cost function “*should describe the desired properties of the sound field reproduced in the control regions as well as the capability of the reproduction system*” [94] (the choice of cost function is re-addressed in detail in Sec. 2.3.3).

To the best of the author’s knowledge, most of the cost functions to optimise <sup>6</sup> are based on the resulting pressure in the zones and the resulting filters. However, some authors have also proposed to jointly optimise the sound pressure and particle velocity [95].

In the literature, usually the CO is not found as a classification on its own (Ref. [33] as a notable exception). Most authors include the different CO among the methods reviewed above. For example, according to which previous technique they were derived from (beamforming or sound field synthesis) [96], according to their general principle of operation control (energy control or sound field synthesis) [86], etc. Although these classifications are accurate, in the author’s opinion, the main difficulty lays in the fact that different CO methods -or combinations- can be used to solve the same problem, *e.g.*, in Ref. [97] the authors worked with a “combined cost function” (energy control and “multipoint” SFS) to solve a superdirective beamforming problem (this is re-addressed in detail later).

Moreover, depending on the distribution of sources, the distribution of the zones and the cost-function chosen, one can obtain an enormous range of possibilities, *e.g.*, phase-shift like directivity control [34, 98], super-directive beamforming [20, 99, 100, 101, 102, 103], end-fire array [104], multipoint SFS [7, 96, 105], etc.

In the literature one can find numerous examples employing CO to solve more “classical” problems. For example in Fig. 2.11 one can find three implementations of the CO methods using three different outlines of sources and zones for three applications.

---

<sup>6</sup>In this thesis optimisation is considered either maximisation or minimisation of the cost function depending on the problem.

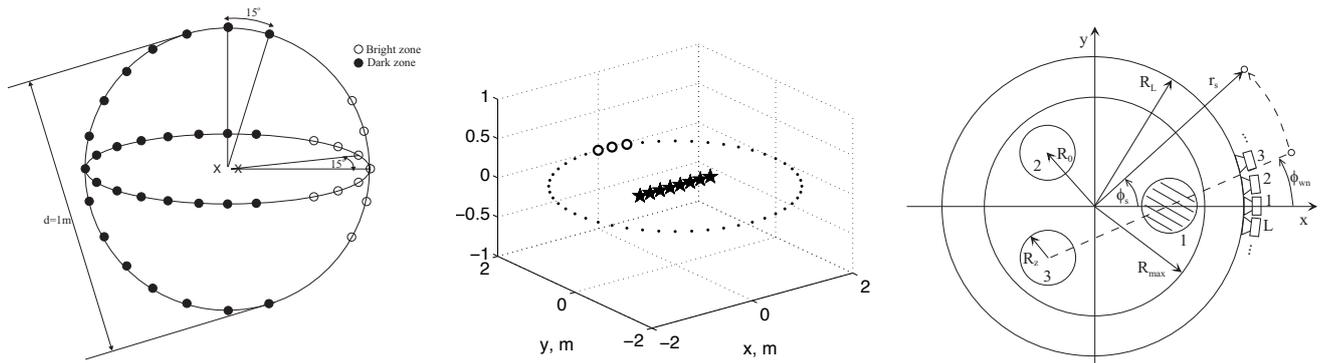


Figure 2.11: CO examples of array and zones distributions for different applications: directivity control of a two source array [34] (left), super-directive beamforming using an eight source array [11] (centre) and multizone SFS [7] (right).

On the left, a two-source array where the control points are placed at a fixed distance of 1 m surrounding the sources (a sort-of 3D directivity control) [34], in the centre, beamforming using an array of eight sources [11] and on the right, multi-zone (multi-point) SFS using circular array of  $L$  sources [7].

The main advantage of the CO methods is probably the flexibility of using the superposition of IRs, which in theory, can be used on any loudspeaker layout [78]. One can solve either more “classical” problems, *e.g.* implementing circular array and synthesizing a plane wave, which results in similar limitations as the analytical methods [78]. Or one can use it to solve problems with more arbitrarily distributed sources and zones and finding solutions for those problems that might not be simple or possible to find analytically.

Furthermore, these methods can be either applied using closed-box sources or improve their performance with cheap passive directive-sources like the “resistance-box” source [16, 99]. Moreover, if the IRs used for the calculation of the filter combination are acquired “in-situ” they would also contain the contributions of the environment as well as any variability due to manufacturing. Another advantage of the CO methods is that it is possible to include in the cost-function the capabilities of the reproduction system as part of the design [94].

Nevertheless, the fact that a solution is found regardless of the source-zones-environment combination, does not mean neither that the solution found meets the requirements nor that it is feasible. In addition, due to the large field of applications, as well as the number of variables involved in the definition of the systems makes it difficult to extrapolate/compare among different works [94, 106].

It might not be so obvious why in this thesis the CO methods are considered as a design criteria on its own, because they can be simply taken as an alternative procedure to calculate the driving signals of any of the multi-transducer methods analysed before (for example the comparisons found in Refs. [87, 88] between analytical SFS and CO methods). It has been highlighted by Møller in 2019 [94], that the system’s layout<sup>7</sup> has a great influence on the resulting performance. Therefore, in the author’s opinion, the optimisation not only comprises the numerical optimisation but also the definition of the geometric layout of the system which might not necessarily coincide with the more classical ones (circular arrays, spherical arrays, linear arrays, etc.).

Some authors proposed to incorporate the source distribution in the design of the system. For example, in 2018 Sanalattii [107] employed the CO methods to design a compact-directive multi-loudspeaker source for acoustic measurements. In order to find the best loudspeaker distribution, the author used the radiation modes method to analyse the vibrational deformation of the geometry and position the loudspeakers according to the radiation characteristics of the source. An example closer to the application of this thesis can be found in Ref. [19], where the authors designed a selection process of sources located in the ceiling of a car cabin, in order to generate PSZ between the front a rear seats.

<sup>7</sup>Distribution both control regions and the corresponding acoustic sources.

Although there are as many advantages to use CO as challenges, these are the main -if not the only- candidates to fulfil the system's constraints. Lastly, it is important to mention that there is a considerable amount of publications involving the generation of PSZ using CO method, particularly implemented in car cabins, *e.g.*, [14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24].

## 2.2.4 Conclusions

Throughout this section, one reviewed the methods that can be -and in many cases have been- used for the generation of PSZ.

This summary of methods could have been tackled from many different angles: according to which previous technique the methods were derived from (beamforming or sound field synthesis) [96], according to their general principle of operation control (energy control or sound field synthesis) [86], according to the frequency band of operation [1], according to the room influence models<sup>8</sup> [94], according to the number of transducers needed for their application, etc.

In this thesis, one chose to classify the methods according to their criteria used for their design, the directivity control, the synthesis of a pressure field or the optimisation of a cost function. All the methods reviewed in this thesis are synthesized in Fig. 2.12.

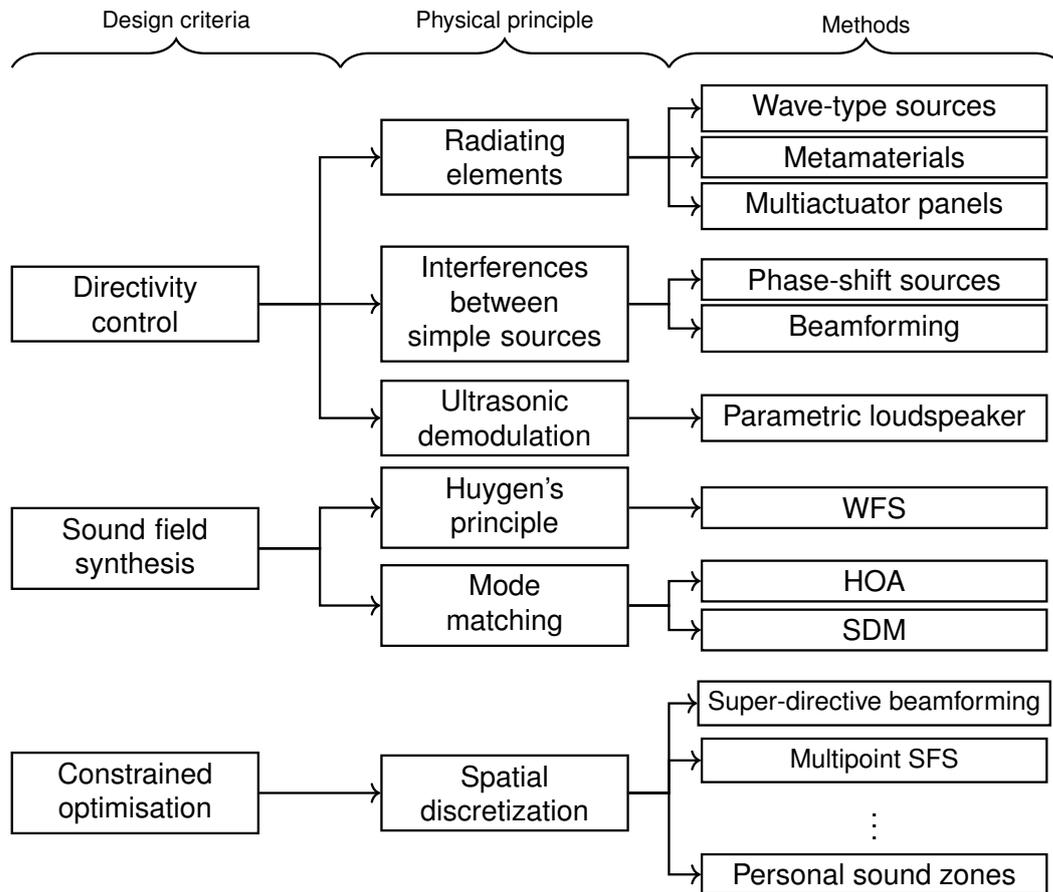


Figure 2.12: Summary of methods.

This type of classification was chosen in order to illustrate that the different methods inside the CO family depend on their application -or the resemblance to another method. All the different CO methods aim

<sup>8</sup>From Moller 2019 [94]: Expression of the sound field in the control regions which results from a specific input signal to the speakers

towards the optimisation of a cost function and their physical principle is always the spatial discretisation of the physical problem.

Lastly, in this thesis the CO methods are chosen as a full design criteria towards the final application. Furthermore, in this thesis the CO is addressed as a threefold problem: the choice of cost function that better-represents the problem in hand, the definition of the “zones” or points to control and the definition of the sources (characteristics and distribution) according to the previous two definitions and the acoustic environment.

In the following section one is going to continue with the literature review but going deeper into posing the PSZ problem as a CO.

## 2.3 PSZ as a constrained optimisation problem

Two different ways can be found in the PSZ literature to pose this filter-optimisation problem: a deterministic approach or an adaptive (iterative) approach.

In this chapter the focus is put on the deterministic approach because is the most-commonly used in the PSZ whereas the adaptive approach is generally used in other applications like transaural sound reproduction or active noise control [108]. Nevertheless, the adaptive approach presents some attractive not so widely-studied characteristics that will be addressed in Chapter 5.

### 2.3.1 Deterministic approach: expressing the problem in the time and frequency domains

For more than 15 years -since 2002 [2]- the calculation of the PSZ filters has been carried out through a deterministic approach (with some notable exceptions *e.g.* Refs. [13, 109, 110, 111]).

The deterministic approach assumes:

- The IRs  $h_{k,l}$  fully/perfectly parametrise the zones and they are time invariant (at least if high performance is desired [13]).
- One knows the characteristics of the input signal, usually defined as a spectrally-flat signal like white Gaussian noise or a Dirac delta function ( $\delta[n]$ ). For simplicity, in this chapter a  $\delta[n]$  is used (as in [26, 97, 106, 112, 113]), but in some cases it could be possible to generalise the formulation for different input signals [97].

Therefore, in order to be able to pose the cost function and its physical implications, the two most widely used formulations of the deterministic approach are reviewed: the time domain and the frequency domain. Nevertheless, it is also possible to use the polynomial matrix approach using the backward shift operator as in Ref. [21, 114]. From this point forward, the sub-index  $s$  is ignored because the problem is reviewed for only one signal at the time.

#### Time domain

Firstly, the time domain formulation is introduced following the works of [26, 97, 106, 112, 113] among others. As it is usually done in the literature, to simplify Eq. 2.3 it is assumed that the input signal is a  $\delta[n]$ , therefore the pressure at microphone  $k$  can be expressed as

$$p_k[n] = \sum_{l=1}^L \sum_{m=0}^{M-1} h_{k,l}[m] w_l[n-m]. \quad (2.4)$$

Furthermore, from Eq. 2.4, the sampled pressure of length  $(M + I - 1 \times 1)$  at microphone  $k$  described by the sum of  $L$  impulse responses can be expressed using a matrix notation as

$$\mathbf{p}_k = \mathbf{H}_k \mathbf{w}, \quad (2.5)$$

being  $\mathbf{w} = [\mathbf{w}_1^T, \dots, \mathbf{w}_L^T]^T$  the  $(ML \times 1)$  vector of concatenated filter's IRs and

$$\mathbf{H}_k = [\mathbf{H}_{k1}, \mathbf{H}_{k2}, \dots, \mathbf{H}_{kL}], \quad (2.6)$$

where  $\mathbf{H}_{kl}$  the  $(M + I - 1 \times M)$  convolution matrix defined as

$$\mathbf{H}_{kl} = \begin{bmatrix} h_{k,l}[0] & 0 & 0 \\ \vdots & \ddots & \vdots \\ h_{k,l}[I-1] & \ddots & 0 \\ 0 & \ddots & h_{k,l}[0] \\ \vdots & \ddots & \vdots \\ 0 & 0 & h_{k,l}[I-1] \end{bmatrix}. \quad (2.7)$$

Therefore, to express the complete problem -for the  $K$  microphones- it is possible to arrange Eq. 2.7 into a  $(K(M + I - 1) \times ML)$  matrix as

$$\mathbf{H} = [\mathbf{H}_1^T, \mathbf{H}_2^T, \dots, \mathbf{H}_K^T]^T. \quad (2.8)$$

Hence, using Eq. (2.8), the vector of sampled IRs for the  $K$  microphones can be written as

$$\mathbf{p} = \mathbf{H}\mathbf{w} \quad (2.9)$$

being  $\mathbf{p} = [\mathbf{p}_1^T, \mathbf{p}_2^T, \dots, \mathbf{p}_K^T]^T$ .

It is important to highlight that, the fact that this whole formulation is based on the system's IRs, Eq. 2.9 contains the information of the complete system, *i.e.* as sampled signals are being used, it describes the complete bandwidth from 0 Hz up to the Nyquist frequency.

### Frequency domain

Analogously to the time domain, the deterministic approach can be also posed in the frequency domain. This formulation was firstly used in PSZ by [2] and it has been the most widely used ever since. The main difference with the time domain formulation is that it represents the overall-problem as a function of the frequency. Therefore, from Eq. 2.3, assuming a flat-spectrum input signal, for a single angular frequency  $\omega$ , the pressure at microphone  $k$  can be expressed in the frequency domain as

$$\hat{p}_k(\omega) = \sum_{l=1}^L \hat{g}_{k,l}(\omega) \hat{q}_l(\omega), \quad (2.10)$$

being  $\hat{q}_l(\omega)$  the complex weight applied to loudspeaker  $l$ ,  $\hat{g}_{k,l}(\omega)$  the complex transfer function (TF) between microphone  $k$  and loudspeaker  $l$  and  $(\hat{\ })$  denoting the frequency domain.

Moreover, from Eq. 2.10 it is possible to express the full problem for the  $K$  microphones (for  $\omega$ ) as

$$\hat{\mathbf{p}}(\omega) = \hat{\mathbf{G}}(\omega) \hat{\mathbf{q}}(\omega), \quad (2.11)$$

being

$$\hat{\mathbf{q}}(\omega) = [\hat{q}_1(\omega), \hat{q}_2(\omega), \dots, \hat{q}_L(\omega)]^T \quad (2.12)$$

and

$$\hat{\mathbf{G}}(\omega) = \begin{bmatrix} \hat{g}_{1,1}(\omega) & \hat{g}_{1,2}(\omega) & \dots & \hat{g}_{1,L}(\omega) \\ \hat{g}_{2,1}(\omega) & \hat{g}_{2,2}(\omega) & \dots & \hat{g}_{2,L}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{g}_{K,1}(\omega) & \hat{g}_{K,2}(\omega) & \dots & \hat{g}_{K,L}(\omega) \end{bmatrix} \quad (2.13)$$

the  $(K \times L)$  matrix of complex transfer functions.

Unlike the expression obtained for the time-domain formulation (Eq. 2.9), the expression of the pressure obtained in Eq. 2.11 contains the information of the full problem ( $K$  microphones and  $L$  sources) but for a single angular frequency  $\omega$ .

### 2.3.2 Metrics

Before attempting to calculate the optimal filters, it is important to define the metrics that permit to evaluate the performance of the system according to the needs of the studied application. Therefore, four metrics are chosen from the literature, each of which is in charge of overseeing a different constraint. One metric to evaluate the “separation” between the  $BZ$  and  $DZ$ , one to evaluate the “transparency” of the system, one to evaluate the power consumption of the system and one to evaluate if there might be some numerical problems.

It is important to highlight that these metrics were chosen because they allow to tackle this particular application, however if the aim of the system was other, one can find in the PSZ literature different metrics. For example, if one was aiming towards a plane-wave reproduction in the  $BZ$  it might be useful to add a metric that measures the extent to which the sound field resembles a plane wave [96].

In order to evaluate the frequency-dependence of the system’s performance, the following metrics are developed using the frequency domain notation explained in Sec. 2.3.1, for compactness, the  $\omega$  is ignored.

#### Acoustic Contrast

In 2002, Choi et.al [2] stated that the acoustic potential energy ( $\epsilon$ ) determines the magnitude of sound perception. Therefore, to measure the “separation” between the zones it is possible to compare the acoustic potential energy between the two desired zones; i.e. the Acoustic Contrast ( $AC$ ) between the zone defined as  $BZ$  and the zone defined as  $DZ$ , as in

$$AC = \frac{\hat{\epsilon}_{BZ}}{\hat{\epsilon}_{DZ}}, \quad (2.14)$$

being  $\hat{\epsilon}_{BZ}$  and  $\hat{\epsilon}_{DZ}$  the potential energy in the  $BZ$  and in  $DZ$  respectively. The potential energy can be approximated by the sum of the potential energy density at a number of discrete locations which is proportional to the sum of the squared pressures [3] (normalised by the number of control points, in our application microphones). Therefore, using the notation introduced in Sec. 2.3.1 it can be defined as:

$$\hat{\epsilon}_{BZ} = \frac{1}{K_{BZ}} \hat{\mathbf{q}}^H \hat{\mathbf{G}}_{BZ}^H \hat{\mathbf{G}}_{BZ} \hat{\mathbf{q}} \quad \text{and} \quad \hat{\epsilon}_{DZ} = \frac{1}{K_{DZ}} \hat{\mathbf{q}}^H \hat{\mathbf{G}}_{DZ}^H \hat{\mathbf{G}}_{DZ} \hat{\mathbf{q}}. \quad (2.15)$$

being  $K_{BZ}$  and  $K_{DZ}$  the number of microphones used to sample the  $BZ$  and  $DZ$  respectively.

Consequently, the  $AC$  from Eq. 2.14 becomes

$$AC = \frac{K_{DZ} \hat{\mathbf{q}}^H \hat{\mathbf{G}}_{BZ}^H \hat{\mathbf{G}}_{BZ} \hat{\mathbf{q}}}{K_{BZ} \hat{\mathbf{q}}^H \hat{\mathbf{G}}_{DZ}^H \hat{\mathbf{G}}_{DZ} \hat{\mathbf{q}}}. \quad (2.16)$$

Usually in the literature, the  $AC$  is expressed in dB and the two controlled zones have the same dimensions and are equally-sampled, therefore the  $AC$  is also found without the  $\frac{K_{DZ}}{K_{BZ}}$  term.

Additionally, throughout this thesis the Acoustic Contrast Gain ( $AC_G$ ) it is also employed as the difference in dB between the resulting  $AC$  of two different systems.

#### Effort

The effort ( $E_{ff}$ ) is the energy that the loudspeaker system requires to perform the PSZ, relative to the energy required for a reference (source/s, monopole/s) to produce the same pressure level in the bright zone. It is defined as

$$E_{ff} = 10 \log\left(\frac{\hat{\mathbf{q}}^H \hat{\mathbf{q}}}{|\hat{q}_r|^2}\right), \quad (2.17)$$

where  $\hat{q}_r$  is the weight that one has to apply to the reference to achieve the same level in the  $BZ$ .

An effort of  $E_{ff} = 0$  dB means that the array requires the same energy as the reference source to generate the PSZ,  $E_{ff} < 0$  means an improvement in the efficiency of the system and an  $E_{ff} > 0$  means that the system is requiring more energy [115]. The effort can also be understood as a “loudspeaker power consumption” metric [33] that gives a indication of the current demanded by the array, sound leakage outside the zones and will also be used often as a constraint for the optimisation algorithms.

### Objective: RMS Error

There are several methods that attempt to obtain a desired pressure field in the  $BZ$  ( $\hat{\mathbf{p}}_{dBZ}$ ). It is possible to measure how far from this target field one is by measuring the normalised error  $E_{rr}$  [19]. Therefore, the normalised error ( $E_{rr}$ ) can be expressed as

$$E_{rr} = \frac{(\hat{\mathbf{p}}_{dBZ} - \hat{\mathbf{p}}_{BZ})^H (\hat{\mathbf{p}}_{dBZ} - \hat{\mathbf{p}}_{BZ})}{\hat{\mathbf{p}}_{dBZ}^H \hat{\mathbf{p}}_{dBZ}}. \quad (2.18)$$

Moreover, the expression of the Error can be used to analyse different “target” parameters. In this report, as the aim is to achieve a transparent system, the  $\hat{\mathbf{p}}_{dBZ}$  is defined as the natural response of the system, and the  $E_{rr}$  is used as a metric to evaluate the deterioration (this is re-addressed in Sec. 3.1).

### Condition Number

In computing problems (such as optimisation) ill-conditioning may describe any computation whose output values are very sensitive to small changes in the input data [14]. In the case of having to perform matrix inversions, it is necessary to be sure that the matrix it is not ill-conditioned. The normally-used condition number of a symmetric matrix with respect to its inverse -as it will be the case in this thesis- is given by the ratio of its largest to its smallest eigenvalue [116].

A large condition number indicates that the matrix is close to singular, i.e. the matrix is ill-conditioned [14].

### 2.3.3 The cost function

Now that the pressure in the zones of the complete problem has been formulated (Sec. 2.3.1) and the metrics to oversee the performance have been chosen (Sec. 2.3.2), one can move to the study of the filter calculation. Firstly, it is necessary to pose the cost function to optimise. For this purpose, it is possible to use the Lagrange multipliers method [2].

#### Lagrange multipliers method in a nutshell:

The aim is to solve an optimisation (maximisation or minimisation) under a specific number of constraints [117]. It can be summarised to the following “three-step recipe”:

To obtain the local “maxima” or “minima” of a  $N$  variables function  $R(x_1, x_2, \dots, x_N)$ , subjected to  $J$  constraints  $B_1(x_1, x_2, \dots, x_N) = b_1, B_2(x_1, x_2, \dots, x_N) = b_2, \dots, B_J(x_1, x_2, \dots, x_N) = b_J$  one needs to:

- 1) Express the Lagrangian equation ( $\mathcal{L}$ ) for  $N$  variables and  $J$  constraints as

$$\mathcal{L}(x_1, x_2, \dots, x_N, \Lambda_1, \Lambda_2, \dots, \Lambda_J) = R(x_1, x_2, \dots, x_N) - \sum_{j=1}^J \Lambda_j (B_j(x_1, x_2, \dots, x_N) - b_j). \quad (2.19)$$

being  $\Lambda_j$  the Lagrange multipliers<sup>9</sup>.

---

<sup>9</sup>The Greek letter  $\lambda_j$  is usually used to denote the Lagrange multipliers, but in this thesis it corresponds to the wave-length, hence the use of  $\Lambda_j$ .

2) Apply the gradient of the Lagrangian ( $\nabla\mathcal{L}$ ) and equates it to zero

$$\begin{bmatrix} \frac{\partial\mathcal{L}}{\partial x_1} \\ \vdots \\ \frac{\partial\mathcal{L}}{\partial x_N} \\ \frac{\partial\mathcal{L}}{\partial\Lambda_1} \\ \vdots \\ \frac{\partial\mathcal{L}}{\partial\Lambda_J} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{\partial R}{\partial x_1} - \sum_{i=1}^I \Lambda_i \left(\frac{\partial B_i}{\partial x_1}\right) = 0 \\ \vdots \\ \frac{\partial R}{\partial x_N} - \sum_{i=1}^I \Lambda_i \left(\frac{\partial B_i}{\partial x_N}\right) = 0 \\ B_1(x_1, \dots, x_N) = b_1 \\ \vdots \\ B_J(x_1, \dots, x_N) = b_J \end{bmatrix} \quad (2.20)$$

3) Find a way to solve (2.20) to obtain the critical points (minima and maxima) of  $R$ .

On the one hand, steps 2) and 3) are mostly algebraic manipulation; depending on how the problem is posed -which function under which constraints- it will be easier, harder or even possible to find the analytical expression of the gradient and subsequently its corresponding solution.

On the other hand, step 1) concentrates all the “physical-meaning” of the problem, the cost function  $\mathcal{L}$  is an expression of the physical problem wanted to be optimised. Moreover, many of the performance differences between the different PSZ publications can be explained by the choice of the cost-function.

Therefore, it is necessary to define the terms to build the Lagrange function (Eq. 2.19) that directly tackles the problem in hand, *i.e.* “translates” the physical requirements of the car-problem into functions and constraints having as variables the filters (either  $\mathbf{w}$  or  $\hat{\mathbf{q}}$ ). The challenge lays when establishing:

- What is the best function to minimise/maximise in order to achieve a  $BZ$  and a  $DZ$ ?
- Which constraints need to be added in order to make the system feasible?

In other words, which cost-function should describes the desired properties of the sound field reproduced in the controlled zones as well as the capabilities of the system? [94]

### 2.3.3.1 Lagrangian terms

Consequently, a summary of the different methods -based on the approach described in Sec. 2.3.1- are reviewed in Table 2.1 with their corresponding terms either controlling the  $BZ$  or the  $DZ$ . All the terms present in Table 2.1 can be expressed using the notation introduced in Sec. 2.3.1.

It is necessary to mention that the following analysis is focused on the terms that control the zones, and not other terms that might be added to the cost-function to ensure the feasibility of the system (like effort [116] or an envelope constraint on the resulting filters [118], etc.). The use of these “feasibility” terms is discussed in Sec. 2.3.3.4.

Methods		$\mathcal{L}$ Terms									
		Potential Energy $BZ$	Error $BZ$	Potential Energy $DZ$	Error $DZ$	P. Energy Diff $BZ-DZ$	Variance in the $BZ$	Resp. Variation $BZ$	Resp. Differential $BZ$	Resp. trend estimation $BZ$	Planarity Control $BZ$
Frequency Domain	Brightness Control [2]	↑↑	–	–	–	–	–	–	–	–	–
	Acoustic Contrast Control [2]	✓	–	↓	–	–	–	–	–	–	–
	Least-squares Pressure matching [7]	–	↓	↓	↓	–	–	–	–	–	–
	Acoustic energy difference [119]	–	–	–	–	↑	–	–	–	–	–
	Planarity control [120]	–	–	↓	–	–	–	–	–	–	✓
	Acoustic contrast with limited sound differences [17]	✓	–	↓	–	–	↓	–	–	–	–
	Sound field reproduction with acoustic contrast constraint [121]	✓	↓	↓	–	–	–	–	–	–	–
Time Domain	Broad-band acoustic contrast control [116] (BACC)	✓	–	↓	–	–	–	–	–	–	–
	BACC with response variation [122]	✓	–	↓	–	–	–	✓	–	–	–
	BACC with response differential [112]	✓	–	↓	–	–	–	–	✓	–	–
	BACC with response trend estimation [113]	✓	–	↓	–	–	–	–	–	✓	–
	BACC with pressure matching [97] (BACCPM)	–	↓	↓	↓	–	–	–	–	–	–

Table 2.1: Summary of deterministic methods used for PSZ and their corresponding lagrangian terms controlling the  $BZ$  and the  $DZ$ . (↑) Maximise, (↓) Minimise, (✓) applied penalty/constraint

The PSZ system analysed in this thesis consists in controlling two zones ( $Z_1$  and  $Z_2$ ) simultaneously so, the cost function to optimise should have at the minimum two terms, one for each zone -either as the function to minimise or as a constraint (see Eq. 2.19). For example, the brightness control proposed in Ref. [2] only controls the  $BZ$  and does not consider the  $DZ$ . This method is included in Table 2.1 as an example, other methods only controlling the  $BZ$  and not considering the  $DZ$  in their cost functions can be found in Ref. [28].

At first glance, in Table 2.1 it is possible to observe that when it comes to the  $DZ$ , there is a sort-of agreement between the authors. As it is not possible to obtain complete silence (0 dB SPL), the term handling the  $DZ$  is almost always a potential energy minimisation. This is due to the fact that the potential energy determines the magnitude of a sound perception [2] and it can be defined for the  $DZ$  in the frequency domain as in Eq. 2.15, or in the time domain as

$$\epsilon_{DZ} = \frac{1}{V_{DZ}} \mathbf{w}^T \mathbf{H}_{DZ}^T \mathbf{H}_{DZ} \mathbf{w} \quad (2.21)$$

and analogously it can be also defined in the  $BZ$ .

On the other hand, the discrepancies among the authors are on what to do in the  $BZ$ . For example, if one completely disregards the quality of the reproduction in the  $BZ$  it is possible to add a potential energy constraint in the  $BZ$  and obtain the largest  $AC$  that the system can produce (becoming the well-known “acoustic contrast control” method or ACC) [2], if a planar sound field is wanted in the  $BZ$  one could add a planarity constraint as studied by Coleman *et al.* and directly tackle this requirement (referred to as the planarity control method) [120], if a flat frequency response is sought in the  $BZ$  adding the response trend estimation term introduced by Schellekens *et al.* might be the best way to go (becoming the Broadband acoustic contrast control with a response trend estimation term method) [113], etc.

There is another “philosophy” among the PSZ methods -that usually appears under the SFS methods [86, 96]- based on estimating an error from a defined desired pressure in the zones, also known as the pressure matching (PM) or least-squares method (LSM). This way of looking at the reproduction problem is well known for synthesis of plane waves [123] as well as for adaptive filtering [124] and it was first introduced in the PSZ, specifically for multizone surround sound systems -to the best of the author’s knowledge- by Poletti in 2008 [7] and widely used ever since.

Moreover, if one defines the vector of the desired pressures to obtain in the zones (on every microphone), for example in the time domain

$$\mathbf{p}_d = [\mathbf{p}_{d1}^T, \mathbf{p}_{d1}^T, \dots, \mathbf{p}_{dK}^T]^T, \quad (2.22)$$

being  $\mathbf{p}_{dk}$  the desired sampled pressure of length  $(M + I - 1 \times 1)$  at microphone  $k$ . And analogously for the frequency domain

$$\hat{\mathbf{p}}_d(\omega) = [\hat{p}_{d1}(\omega), \hat{p}_{d2}(\omega), \dots, \hat{p}_{dK}(\omega)]^T, \quad (2.23)$$

being  $\hat{p}_{dk}(\omega)$  the desired complex pressure at microphone  $k$  for the angular frequency  $\omega$ . Hence, it is possible to define the error vector as the difference between these desired pressures and the actual reproduced pressure in the time domain

$$\begin{aligned} \mathcal{E} &= \mathbf{p}_d - \mathbf{p} \\ &= \mathbf{p}_d - \mathbf{H} \mathbf{w}_s. \end{aligned} \quad (2.24)$$

or in the frequency domain, for a single frequency  $\omega$

$$\begin{aligned} \hat{\mathbf{e}}(\omega) &= \hat{\mathbf{p}}_d(\omega) - \hat{\mathbf{p}}(\omega) \\ &= \hat{\mathbf{p}}_d(\omega) - \hat{\mathbf{G}}(\omega) \hat{\mathbf{q}}(\omega). \end{aligned} \quad (2.25)$$

Lastly, the error is usually aimed to be minimised in a least-squares sense, *i.e.* to minimise the sum of the error squares. Therefore, from Eq. 2.24, the least-squares error in the time domain can be expressed from Eq. 2.24 as

$$|\mathcal{E}|^2 = (\mathbf{p}_d - \mathbf{p})^T (\mathbf{p}_d - \mathbf{p}) \quad (2.26)$$

and for the frequency domain from Eq. 2.25 as

$$|\hat{\mathbf{e}}(\omega)|^2 = (\hat{\mathbf{p}}_d(\omega) - \hat{\mathbf{p}}(\omega))^H (\hat{\mathbf{p}}_d(\omega) - \hat{\mathbf{p}}(\omega)). \quad (2.27)$$

Moreover, if one splits Eq. 2.26 between the microphones sampling the  $BZ$  and the ones sampling the  $DZ$  and uses the notation introduced in Sec .2.3.1

$$|\mathcal{E}|^2 = (\mathbf{p}_{dBZ} - \mathbf{H}_{BZ}\mathbf{w})^T (\mathbf{p}_{dBZ} - \mathbf{H}_{BZ}\mathbf{w}) + (\mathbf{p}_{dDZ} - \mathbf{H}_{DZ}\mathbf{w})^T (\mathbf{p}_{dDZ} - \mathbf{H}_{DZ}\mathbf{w}) \quad (2.28)$$

and analogously in the frequency domain from Eq. 2.27, for a single frequency (ignoring the  $\omega$  for compactness)

$$|\hat{\mathbf{e}}|^2 = (\hat{\mathbf{p}}_{dBZ} - \hat{\mathbf{G}}_{BZ}\hat{\mathbf{q}})^H (\hat{\mathbf{p}}_{dBZ} - \hat{\mathbf{G}}_{BZ}\hat{\mathbf{q}}) + (\hat{\mathbf{p}}_{dDZ} - \hat{\mathbf{G}}_{DZ}\hat{\mathbf{q}})^H (\hat{\mathbf{p}}_{dDZ} - \hat{\mathbf{G}}_{DZ}\hat{\mathbf{q}}) \quad (2.29)$$

It is possible to observe that both in Eq. 2.28 and Eq. 2.29 it is possible to retrieve the two terms one controlling the  $BZ$  and one controlling the  $DZ$ .

### 2.3.3.2 Definition of the desired

It is necessary to define the desired pressures in such a way that the minimisation of Eqs. 2.28 and 2.29 tackles the already mentioned constraints. In theory, the desired pressure can be arbitrarily chosen but, it is necessary to have in mind that the definition of the target pressure affects the performance of the overall system (this is re-addressed in detail in Chapter 3). This important remark was studied for example by Olivieri *et al.* in 2015 [102] or in 2017 by Liao *et al.*[19]. The latter, proposed a custom target pressure distribution (as a function of the frequency) to achieve a higher level of acoustic contrast performance.

As it is usually done for PSZ [15, 19, 102, 125], the desired IRs in the  $DZ$  ( $\mathbf{p}_{dZ_2}$ ) is defined in this thesis as a vector of zeros ( $\mathbf{0}$ ), hence from Eq. 2.28

$$|\mathcal{E}|^2 = (\mathbf{p}_{dBZ} - \mathbf{H}_{BZ}\mathbf{w})^T (\mathbf{p}_{dBZ} - \mathbf{H}_{BZ}\mathbf{w}) + \mathbf{w}^T \mathbf{H}_{DZ}^T \mathbf{H}_{DZ} \mathbf{w} \quad (2.30)$$

and in the frequency domain from Eq. 2.29

$$|\hat{\mathbf{e}}|^2 = (\hat{\mathbf{p}}_{dBZ} - \hat{\mathbf{G}}_{BZ}\hat{\mathbf{q}})^H (\hat{\mathbf{p}}_{dBZ} - \hat{\mathbf{G}}_{BZ}\hat{\mathbf{q}}) + \hat{\mathbf{q}}^H \hat{\mathbf{G}}_{DZ}^H \hat{\mathbf{G}}_{DZ} \hat{\mathbf{q}}. \quad (2.31)$$

In both domains (Eq. 2.30 and Eq. 2.31) the  $DZ$  terms become a potential energy term (proportional to Eq. 2.21 and Eq. 2.15). This particular case of the error minimisation algorithms is found in the literature in the frequency domain as the ‘‘combined solution’’ (CS) [126] or in the time domain as the broad-band acoustic contrast control with pressure matching penalty (BACCPM) [97]. It is important to highlight that in the literature the CS and the BACCPM are found with a weighting factor between the two terms (this is re-addressed below).

On the other hand, the desired pressure in the  $BZ$  needs to also be defined in agreement with the problem’s constraints. Due to the fact that the aim is to design a transparent system, the desired pressures in the  $BZ$  are defined as the pressures when the system is being used without any processing (this is re-addressed when defining the system in Sec. 3.1).

For this study, both Eq. 2.30 and Eq. 2.31 are chosen as possible Lagrangian -cost- functions to solve the PSZ problem.

### 2.3.3.3 The filters

Hence, now that the cost function is chosen, one can follow the Lagrange multipliers procedure ( $\nabla \mathcal{L} = \mathbf{0}$ ) to calculate the filters. Hence, assuming that  $L \leq K$  (an overdetermined system [7, 14, 124]), the expression of the filters in the time domain is (from Eq. 2.30)

$$\mathbf{w} = [\mathbf{H}_{BZ}^T \mathbf{H}_{BZ} + \mathbf{H}_{DZ}^T \mathbf{H}_{DZ}]^{-1} \mathbf{H}_{BZ}^T \mathbf{p}_{dBZ} \quad (2.32)$$

and in the complex weights frequency domain (from Eq. 2.31)

$$\hat{\mathbf{q}} = \left[ \hat{\mathbf{G}}_{BZ}^H \hat{\mathbf{G}}_{BZ} + \hat{\mathbf{G}}_{DZ}^H \hat{\mathbf{G}}_{DZ} \right]^{-1} \hat{\mathbf{G}}_{BZ}^H \hat{\mathbf{p}}_{dBZ}. \quad (2.33)$$

It is necessary to remember that the expression of  $\hat{\mathbf{q}}$  in Eq. 2.33 is a vector of complex weights for a single frequency, whilst the expression of  $\mathbf{w}$  in Eq. 2.32 is the concatenation of the resulting filter's IRs. In addition, when comparing the size of the matrices in Table 2.2, it is possible to observe that the frequency domain formulation handles considerably smaller matrices than the time domain.

Frequency Domain	Matrix	$\hat{\mathbf{G}}_{BZ}$	$\hat{\mathbf{G}}_{BZ}^H \hat{\mathbf{G}}_{BZ}$	$\hat{\mathbf{p}}_{dBZ}$	$\hat{\mathbf{q}}$
	Size	$K_{BZ} \times L$	$L \times L$	$K_{BZ} \times 1$	$L \times 1$
Time Domain	Matrix	$\mathbf{H}_{BZ}$	$\mathbf{H}_{BZ}^T \mathbf{H}_{BZ}$	$\mathbf{p}_{dBZ}$	$\mathbf{w}$
	Size	$K_{BZ}(M + I - 1) \times ML$	$ML \times ML$	$K_{BZ}(M + I - 1) \times 1$	$ML \times 1$

Table 2.2: Matrix size comparison for both formulations, using  $L$  loudspeakers,  $K_{BZ}$  microphones in the  $BZ$ ,  $M$  filter coefficients and impulse responses of length  $I$ .

Furthermore, to obtain the matrix  $\hat{\mathbf{G}}_{BZ}^H \hat{\mathbf{G}}_{BZ}$  the number of operations (both sums and products) is equal to  $(2K_{BZ} - 1)L^2$  whilst to obtain  $\mathbf{H}_{BZ}^T \mathbf{H}_{BZ}$  is  $[2K_{BZ}(M + I - 1) - 1](ML)^2$ .

On the one hand, the time domain formulation directly calculates the filter's impulse responses for the complete bandwidth, which results in handling larger matrices and performing more operations to calculate the filters. On the other hand, although the frequency domain handles smaller matrices, its expression considers each frequency as an independent problem. Therefore, it is necessary to calculate the complex weights for several frequencies (over the desired bandwidth to control) and apply some sort-of frequency truncation and an inverse Fourier transform in order to obtain the filter IRs [106, 125] without guaranteeing the causality of the filters (presence of pre-ringing) and most likely requiring an additional delay (modelling delay)[97].

Two works have been recently published, Refs. [27, 127], proposing to separate the PSZ optimisation in sub-bands in order to reduce the computational costs, the size of the filters and increase the versatility when a broad-band system is designed. Following the sub-band filters proposed by Ref. [127], for this thesis, it might only be possible to divide the studied bandwidth in no more than two sub-bands. Nevertheless, this approach should be considered when a full band system is assembled.

The implications of either using the time domain expression (Eq. 2.32) or the frequency domain expression (Eq. 2.33) are thoroughly reviewed on the actual application in Sec. 3.5.

#### 2.3.3.4 Feasibility terms and weighting parameters

In addition to the terms optimising the  $BZ$  and  $DZ$ , there are other terms and control parameters that can be added to the cost functions (Eqs. 2.30 or 2.31). These extra terms are either added to have more control over the existing cost function or to incorporate other constraints to the system.

##### Effort regularisation

The feasibility constraint that has been used in every PSZ formulation is the effort regularisation term also known as Tikhonov regularisation. To the best of the authors knowledge, this regularisation term can be found in every PSZ cost function as a default. For example, all the methods listed in Table 2.1 have been published with an effort constraint.

This constraint/penalty has been widely-used in active control techniques in order to reduce possible ill-conditioned problems that might arise from the inversion of the matrices (see Eqs. 2.32 and 2.33) which has the effect of reducing the sum of the modulus squared source strengths [116] (proportional to the

$E_{ff}$  metric in Sec. 2.3.2). In addition, this term has been also used to provide more stable system in the presence of noisy TFs [128] and inaccurate IRs [129] (the inaccuracies in the TFs are studied in detail in Sec. 4).

Although this constraint had already been used for PSZ applications (*e.g.* Refs. [2, 7]) it was thoroughly reviewed by Elliott *et al.* in Refs. [116, 129].

The effort constraint has been applied in the frequency domain (*e.g.* [116, 14, 96]) in the expression of the complex weights (from Eq. 2.33)

$$\hat{\mathbf{q}} = \left[ \hat{\mathbf{G}}_{BZ}^H \hat{\mathbf{G}}_{BZ} + \hat{\mathbf{G}}_{DZ}^H \hat{\mathbf{G}}_{DZ} + \Lambda \mathbf{I} \right]^{-1} \hat{\mathbf{G}}_{BZ}^H \hat{\mathbf{p}}_{dBZ} \quad (2.34)$$

being  $\mathbf{I}$  the identity matrix and  $\Lambda$  is the Lagrange multiplier -the effort tuning/regularisation parameter. The ‘‘appropriate’’ imposition of the effort constraint is defined by choosing the value of  $\Lambda$ . It can be defined as a constant value, to equally constraint all the frequencies, or a different amount of regularisation per frequency according to some criteria, usually to reduce the array’s effort [102, 129].

On the other hand, the effort constraint has also been applied in the time domain (*e.g.* [97, 106, 118]) resulting in the filter expression (from Eq. 2.32) as

$$\mathbf{w} = \left[ \mathbf{H}_{BZ}^T \mathbf{H}_{BZ} + \mathbf{H}_{DZ}^T \mathbf{H}_{DZ} + \Lambda \mathbf{W} \right]^{-1} \mathbf{H}_{BZ}^T \mathbf{p}_{dBZ} \quad (2.35)$$

being  $\Lambda$  the Lagrange multiplier (the effort tuning parameter) and  $\mathbf{W}$  is the effort weighting matrix as in Ref [97]. If  $\mathbf{W}$  is chosen as a the identity matrix, all the frequencies are equally constrained, but it can be defined to perform a frequency dependent regularisation defining the  $ML \times ML$  matrix

$$\mathbf{W} = \begin{bmatrix} \mathbf{B}^T \mathbf{B} & 0 & \dots & 0 \\ 0 & \mathbf{B}^T \mathbf{B} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{B}^T \mathbf{B} \end{bmatrix} \quad (2.36)$$

being the  $M + I - 1 \times M$  matrix

$$\mathbf{B} = \begin{bmatrix} b(0) & 0 & 0 \\ \vdots & \ddots & \vdots \\ b(M-1) & \ddots & 0 \\ 0 & \ddots & b(0) \\ \vdots & \ddots & \vdots \\ 0 & 0 & b(M-1) \end{bmatrix} \quad (2.37)$$

where  $\mathbf{b}$  is the FIR filter used to generate the frequency dependence. For example, if  $\mathbf{b}$  is a low-pass filter, it imposes the regularisation in the lower frequencies, if a high-pass filter is used, the higher frequencies [97]. This approach is useful assuming that the ‘‘needed’’ regularisation is known beforehand.

Nevertheless, the addition of this regularisation term comes at a cost of a reduction of the performance of the system [129]. Furthermore, if the problem is not ill-conditioned, is not demanding a large effort and the presence of noise is low, theoretically there is no need to add this regularisation parameter and it would be possible to calculate the filters as expressed in Eqs. 2.32 and 2.33.

It was proposed by Gauthier in 2016 [130] to decompose the PSZ problem in a singular value decomposition-based domain to separate both zones using two sets of singular values and three sets of singular vectors. This was later implemented by Zhu in 2019 [131] to formulate regularisation as the modification of the assumed modal capabilities in designing sound zone reproduction. Although this type of modal decomposition of the PSZ shows promising results, its study is out of the scope of this dissertation.

In this thesis, a criteria is chosen in order to only apply the effort constraint if needed, *i.e.* if the condition number grows considerably or if a large effort is demanded to the array. Hence, unless otherwise specified, the regularisation term is ignored in the filter calculation ( $\Lambda = 0$ ).

### Other feasibility terms and weighting parameters

There are other terms that have been added to the cost function, associated with the feasibility of the system more than the physical generation of a  $BZ$  and  $DZ$ . For example, in 2017 Møller *et al.* introduced a penalty term to BACCPM in order to control the shape of the resulting filters [118].

A recent case that needs to be mentioned is the method proposed by Lee *et al.* in 2020 to shape the reproduction errors in a perceptual meaningful way. This is achieved by calculating weighting filters according to the input signal characteristics and the human auditory system (specifically masking curves) and it showed perceptually better performances in spite of not necessarily achieving better physical metrics [132]. This approach is a different application of the frequency depending weighting in the time domain formulations (originally used in the effort regularisation [97]) explained above. Although, it could be possible to introduce this principle to evaluate the improvement in the perceived-performance, this perceptual weighting is ignored from the following development.

On the other hand, it was briefly mentioned that the combined solution's cost functions (both in the frequency domain [126] as well as in the time domain [97]) is employed using a weighting parameter to handle the "importance" between the error minimisation term and the potential energy minimisation term. For example, in the frequency domain (as in Ref. [126]), a weighting parameter  $0 < \kappa < 1$  is added to the cost function in Eq. 2.31 which becomes

$$\hat{\mathcal{L}} = (1 - \kappa)(\hat{\mathbf{p}}_{dBZ} - \hat{\mathbf{G}}_{BZ}\hat{\mathbf{q}})^H(\hat{\mathbf{p}}_{dBZ} - \hat{\mathbf{G}}_{BZ}\mathbf{q}) + \kappa\hat{\mathbf{q}}^H\hat{\mathbf{G}}_{DZ}^H\hat{\mathbf{G}}_{DZ}\hat{\mathbf{q}}. \quad (2.38)$$

If  $\kappa \approx 1$ , the algorithm will give preference to the potential energy minimisation but if  $\kappa \approx 0$  it will give preference to the error minimisation.

It is necessary to clarify that the addition of this weighting parameter in the case of the combined solution does not represent a direct trade-off between the  $AC$  and the  $E_{rr}$ . Moreover, if one looks at  $\kappa$  at extreme values (in Eq. 2.38); when  $\kappa = 1$  the cost function only has one term controlling the  $DZ$  (which results in null-filters  $\hat{\mathbf{q}} = \mathbf{0}$ ), and if  $\kappa = 0$  is only an error minimisation problem in the  $BZ$  without any attenuation in the  $DZ$  [133].

Therefore, if a proper trade-off between  $AC$  and  $E_{rr}$  is sought, it is possible to express the deterministic problem as a trade-off between the BACC and the BACCPM using the method inspired by the variable span line [133, 134].

The approach of expressing the cost function as a weighted sum of the different terms (functions or constraints) can be in theory extended to as many constraints as needed. For example in Ref. [13], the authors expressed the cost function as a five weighted terms equation. However, every term added means an extra constraint as well as an extra weighting parameter to define [94], *i.e.* an extra trade-off incorporated into the problem that can compromise the original cost-functions in charge of generating the zones.

In this thesis the original least-squares minimisation error, *i.e.*, the "classic" PM/BACCPM already fulfils the application's needs and are chosen as the best cost functions without any weighting parameter (or equivalently in Eq. 2.38,  $\kappa = 0.5$ ).

## 2.4 Perceptual evaluation

When designing a PSZ it is necessary to have in mind that, regardless of the technology used, it is unlikely to be able to generate a perfect separation between the zones (pressure in the  $DZ=\mathbf{0}$ ).

Moreover, although this thesis is focused on the physical aspects of the generation of PSZ, other authors have focused their research in the perceptual evaluation of the listening experience. Whilst the physical metrics (see Sec. 2.3.2) are useful to evaluate the physical performance of the system, the success of a system ultimately rests on the experience of the listener [135].

The first challenge of the perceptual evaluation of PSZ was to find an subjective attribute to rate the effect of the interferer audio programme over the target audio programme. The results published in 2014 by Francombe *et al.* [136] showed that the perceptual attribute "distraction", was related to the level of

the interferer and it produced the strongest agreement between subjects. This attribute was described as: “*How much the alternate audio pulls your attention or distracts you from the target audio*”.

In 2015, Baykaner *et al.*[30] conducted an investigation into the relationship between target quality, distraction and the overall quality of the PSZ listening experience employing different algorithms (ACC, PM and different variations of the planarity control [120]). Although the performance levels achieved by the different methods are limited to the conditions tested in this study, the authors showed that the target quality and the distraction had an equivalent importance to when determining the overall quality. Additionally, it was observed that some types of programs (like pop music and sports commentary) were more robust to interference than others (like classical music).

In 2016, Rämö *et al.* [135] evaluated the user’s perceptual distraction caused by interfering audio programmes in a complex PSZ system (68 loudspeakers) and validated the perceptual model of distraction previously implemented by Francombe [137]. A listening test, involving 26 listeners revealed that the distraction is approximately inversely correlated with the target-to-interferer ratio (TIR).<sup>10</sup> Additionally, it was observed that to achieve a distraction level of “0” (*not at all distracting*) it is necessary to achieve a target-to-interferer ratio of approximately 27 dB.

For more information on the perceptual evaluation of PSZ, refer to the doctoral dissertations of Francombe [137] and Baykaner [138].

Although the previous studies provide an idea of the physical requirements of the PSZ system [94], it is necessary to adapt these results to the this dissertation’s requirements. In this dissertation, one is going to aim to achieve an  $AC$  of 30 dB across the studied frequency band, but it has been previously established that around 25 to 29 dB should suffice [94]. Additionally, one is going to analyse the resulting pressure fields between the two zones in order to achieve the same levels of attenuation uniformly distributed in the zones. Furthermore, an “improvement” metric is also included to quantify the  $AC$  gain of the system using the  $AC_G$  metric (see Sec.2.3.2).

Lastly, it is important to mention that some authors also incorporate characteristics of the human auditory system (masking effects) when defining the PSZ processing. For example, introducing a masking signal in the  $DZ$  to reduce the intelligibility of the interferer programme [139], or the already mentioned weighting filters according to the input signal characteristics and the human auditory system (specifically masking curves) proposed by Lee in Ref. [132].

## 2.5 The procedure: synthesis of the method

Now that an expression of the filters is obtained one can introduce the system’s procedure. The full-procedure can be seen as the following list of steps [13, 94]:

---

Deterministic approach steps:

---

- 1) Place  $K$  microphones to sample the zones to control.
  - 2) Measure every loudspeaker/microphone pair impulse response.
  - 3) Calculate the filters (Eq. 2.32 or Eq. 2.33).
  - 4) Remove the microphones.
  - 5) Apply the calculated filters to the signal to reproduce (feed-forward control system [3]).
- 

These steps bring back the original assumption that the IRs fully-parametrise the problem at hand and that the acoustical characteristics of the system are invariant (established in Sec. 2.3.1). Although there are many studies showing that this is far from the truth, firstly it is necessary to solve the PSZ problem

<sup>10</sup> $TIR$ : the difference in equivalent A-weighted sound pressure levels between target and interferer in one zone, calculated using 10 s integration time [94].

under invariant acoustics, the study of the system when the acoustic characteristics change is thoroughly reviewed in Chapters 4 and 5.

On the other hand, the procedure can be also interpreted as an Input / Control / Output problem as in Fig. 2.13.

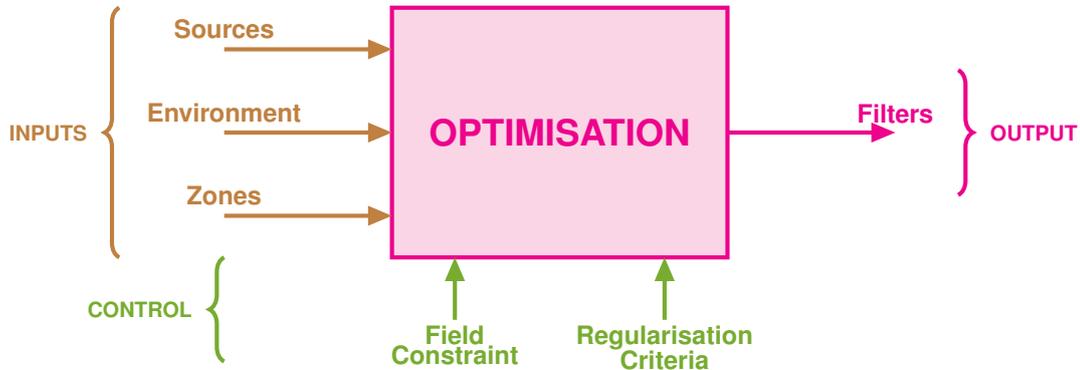


Figure 2.13: Schematic: PSZ problem. Inputs: source, environment and zones; Control: regularisation criteria and field constraint (optional); Output: optimal filter combination.

There are three different elements in Fig. 2.13 that affect the OUTPUT’s calculation -*i.e.* the calculation of the filter combination.

- **OPTIMISATION:** it is the theory developed throughout this chapter, the optimisation criteria chosen to calculate the filters, *i.e.* the chosen cost function.
- **INPUTS:** are the representation of the acoustic conditions which are all condensed in the TF/IR matrices ( $\mathbf{H}$  or  $\hat{\mathbf{G}}$ ). Among them are: the sources (number, type, their location, directivity, etc.), the environment (free-field, reflections, car cabin, diffractive surfaces, etc.) and the zones (size, relative positioning, etc.).
- **CONTROL:** these are elements introduced in the calculation to tune/control how the optimisation reacts to the input parameters; for different optimisation/input combinations different control parameters are needed. For example, the desired pressure, the regularisation criteria, the weighting parameters, etc.

The block diagram in Fig. 2.13 highlights the fact that, although the core of the constrained optimisation methods is which cost function is used for the optimisation, the resulting filters (consequently the resulting performance) also depend on both its inputs as well as its control parameters.

Most authors focus on the development of more and more complicated algorithms, new formulations, new cost functions, etc. but very few (*e.g.* [3, 14, 19, 102, 140]) focus on the repercussion of the inputs and the control parameters. Additionally, one could safely hypothesise that if both the inputs as well as the control parameters are tailored towards the final goal, it would “ease the algorithm’s job”.

Furthermore, in 2016 Moller *et al.* stated: “Commonly, the methods, presented in publications involving sound zones, are implemented under very dissimilar conditions *e.g.* different loudspeaker layouts, listening spaces with different acoustical characteristics, or different selection of bright and dark zone configuration. Furthermore, the exact conditions for the performance evaluation are rarely stated in the literature. It is therefore difficult to compare results between publications due to such differences potentially having a large impact on the evaluated performance of the algorithms” [106]. Namely, due to the number of inputs and control parameters affecting the performance of the system it is necessary to make a specific analysis of each application and be very cautious when taking conclusions from other studies.

Therefore, now that the metrics to evaluate the performance and the method to generate the PSZ are chosen, it is possible to define the inputs and the control elements of the system (see Fig. 2.13). Additionally, one can evaluate which inputs can be tailored, which ones are fixed due to the industrial context of the application, which field constraint is needed and how all this affects the system’s performance.

## 2.6 Summary

In this chapter, several technologies capable of solving the PSZ were reviewed. The methods based on constrained optimisation were chosen because of their potentiality to tackle the requirements and constraints specified by the manufacturer.

These methods are based on defining the cost function that better represent the PSZ application in hand. The use of a least-squares error minimisation was chosen as the most suited cost function, because it allows to define a desired pressure field in the *BZ* whilst doing a potential energy minimisation in the *DZ* [97, 126].

The PSZ problem was posed both in the frequency and in the time domain (Sec. 2.3.1) due to the fact that both formulations tackle the same physical parameters, but in different ways; the time domain for the complete frequency band whilst the frequency domain considers each frequency as an independent problem.

In addition, four metrics were chosen from the literature to evaluate the system's performance (Sec. 2.3.2), one to evaluate the "acoustic-separation" between the zones (*AC*), one to evaluate the degree of transparency of the system ( $E_{rr}$ ), one to evaluate the power consumption of the system ( $E_{ff}$ ) and one to overlook that there are no numerical problems (Cond.). Additionally, a brief review on the studies that were focused on the perceptual evaluation of the PSZ problem was introduced.

The procedure to implement the method was reduced to a series of discrete steps and the overall method analysed as a input/control/output system (Sec. 2.5). Leading the way to the next steps of the research, the study of the input modifications, how do they affect the performance, how they can be tailored (and which) to improve the generation of PSZ in a car cabin.



## Chapter 3

# Invariant acoustics & system definition

As it was reviewed in Chapter 2, even if the best cost function is chosen, the resulting filters depend on the input characteristics *i.e.* the FRFs between the sources and receivers (see Fig. 2.13). Therefore, to get the best performance, it is necessary to design a system anticipating which inputs work better according to the desired effect one wants to achieve. Unfortunately, in a car cabin not every input (sources, environment, zones) can be tailored to maximise performance, *e.g.* the position of the zones is limited to the positioning of the passengers and it is not only difficult but also unlikely to be able to customise the acoustic environment.

In this study, it is assumed that the sources -their location and number- is the only input that can be modified, at least under certain conditions; *e.g.* it is not realistic to think that a source can be installed in the middle of the wind-shield or floating in front of the head of the listener.

It is important to highlight that not every author has tackled the PSZ problem from this angle, many authors have focused their work in the study and development of the algorithm using more “classical” source arrangements. Nevertheless, it is possible to find some studies considering the effect of the sources in the overall performance or simply to reduce costs. For example, in 2014 Coleman *et al.* [140], numerically optimised the choice of the loudspeaker-configuration among 60 loudspeakers uniformly distributed in a circle around the zones.

In the case of cars, the choice of array is usually the already existing door-sources (especially for the lowest part of the spectrum). However, some authors also studied the addition of two-loudspeaker headrests to these door-sources to extend the bandwidth of operation (*e.g.* in Refs. [14, 24] among others). Furthermore, in 2017 by Liao *et al.*, the authors studied the design of a custom array of loudspeakers located in the ceiling of a car cabin [19]. In this dissertation, the ceiling positioning is not considered due to the lack of compatibility with convertible cars and sunroofs.

On the other hand, this section focuses on invariant acoustics, *i.e.* assuming that a full characterisation of the system is available and that this one does not change over time. Hence, it is possible to use a deterministic approach and calculate the best filter combination as defined in Sec. 2.3.3.3.

In order to assess the design of the system, a set of different acoustic environments/conditions (shown in Table 3.1) were chosen. This study is carried out from the simplest acoustic condition (monopoles in free-field) increasing gradually towards the most complex (the car cabin).

Complexity	Environment	Diffractive elements	Type of sources	TF/IR
Nº1	Free-field	-	Monopoles	Simulated (Matlab)
Nº2	Anechoic	Headrest Seat Dummy	Headrest loudspeakers	Measured (Anechoic chamber)
Nº3	Infinite baffles	-	Monopoles Mirror sources	Simulated (Matlab)
Nº4	Anechoic Wooden panels	Headrest Seat Dummy	Headrest loudspeakers	Measured (Anechoic chamber)
Nº5	Car Cabin	Headrest Seat Dummy Car	Headrest loudspeakers	Measured (Peugeot 3008)

Table 3.1: Static acoustics, studied degrees of acoustic

Starting from ideal monopoles in free-field, all the way up to the car cabin (see Table 3.1), each level of acoustic complexity adds information towards the design of a realistic system.

To assist in the evaluation of the different levels of complexity, in the following sections, a frequency domain optimisation is used. This formulation, calculates the optimal weight frequency-by-frequency as independent problems and allows to focus this first study only on the physical limitations. Nevertheless, once an array is defined, an analysis is carried out focusing in the filter's impulse responses, their causality and their effect on the overall performance (in Sec. 3.5).

Lastly, when experimental data are being used (measured IRs), in order to avoid over-determination, two different sets of IRs are used, one to obtain the filters, and a different one to evaluate their performance.

### 3.1 The starting point: system definition

Before attempting any complicated signal processing, it is necessary to choose one system/configuration as a reference to evaluate the system's performance (see Fig. 3.1).

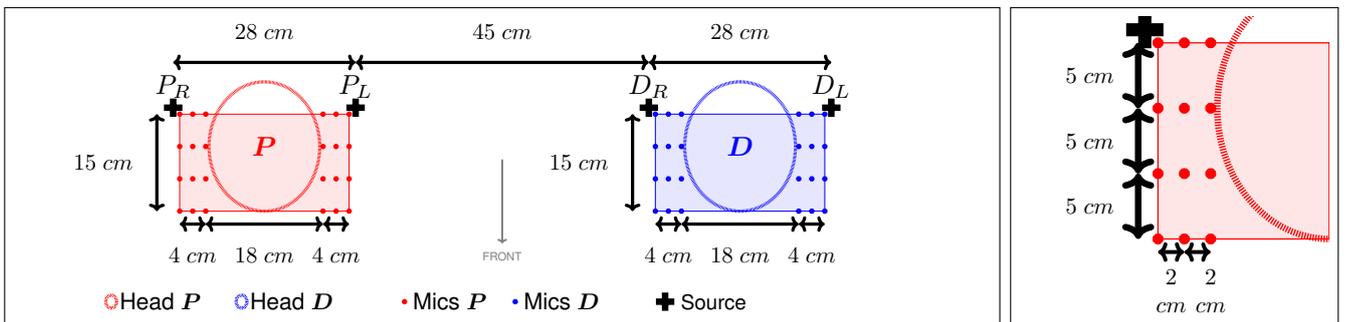


Figure 3.1: Schematic: the starting point, four-source PSZ system.

Firstly, it is necessary to define the zones to control. For this study, it was chosen to control the driver ( $D$ ) and passenger ( $P$ ) areas, *i.e.* two zones separated by 73 cm (between their centres) as in Fig. 3.1. These zones are defined as two rectangular planes of 26 cm  $\times$  15 cm at ears height, in order to represent the physical surface where the heads are located.

Moreover, in the presence of a listener, it is neither possible nor necessary to control the complete zone

(there is no need to control what happens inside the head or on the nose of the listener). Therefore, the zones to control can be reduced to one  $4 \text{ cm} \times 15 \text{ cm}$  rectangle per ear.

Furthermore, each zone is discretised using twelve measuring points per ear with a microphone spacing of  $5 \text{ cm}$  along the  $x$ -axis and  $2 \text{ cm}$  along the  $y$ -axis (see Fig. 3.1); ensuring to have up to  $1.3 \text{ kHz}$  at least 5 points per wavelength. In total, one is using a total of control points  $K = 48$ .

A more user-oriented measuring methodology was proposed by Chang in 2019 [23] using 13 subjects and acquiring the resulting pressure using binaural microphones. The subjects were instructed to sit in two manners, a fixed-posture and a free-posture, and the  $AC$  was calculated from these measurements. For simplicity, this methodology is ignored from this thesis. However, it is important to mention that in all the studied configurations (both simulated and measured) a microphone is always located at the position of the listener’s ears. In addition, in the practical cases not only the performance-metrics are analysed (Sec. 2.3.2) but also the resulting pressure fields in the zones.

Nowadays, it is almost an industrial-standard (especially in cars) to have two loudspeakers in the headrest (see Appendix A) and it has been studied in numerous examples of the PSZ literature (like in [3, 28, 15, 16, 23, 24] among others). On the one hand, the near-field arrays allow to extend the bandwidth of the system and improves the coupling between the microphones and the control sources because their responses are mainly dominated by the direct sound [16]. On the other hand, the main drawback is that the effects of objects in close proximity to the sources may affect the performance [16].

As the “starting point” of this thesis, an array of two sources per seat ( $L = 4$ ) is used separated by  $28 \text{ cm}$  at the same height as the zones. These sources are denoted  $P_R$  and  $P_L$  as the passenger’s right and left sources respectively and analogously for the Driver ( $D_R$  and  $D_L$ ). For simplicity, throughout this thesis it is assumed that the passenger is the  $BZ$  and the driver is the  $DZ$ .

As mentioned in (Sec. 2.3.3), an error minimisation term is used to control the reproduction in the  $BZ$ , so it is necessary to define the target pressure. In order to obtain a transparent performance, the target pressure is defined assuming the reproduction of a mono signal; *i.e.* the resulting pressure field of the two sources the closest to the  $BZ$  ( $P_R$  and  $P_L$  in Fig. 3.1) when driven in phase. In other words,  $P_R$  and  $P_L$  filtered by perfect unit-sample/Dirac delta function each ( $\delta[n]$ ). By defining the target pressure in this way one also ensures that the system is capable of reproducing the demanded pressure field. However, if an arbitrary target pressure is used, one might need to evaluate the capability of the system to reproduce it (as studied in Ref. [105] for example).

Moreover, this definition of the target pressure allows to use the  $E_{rr}$  as an indicator of the transparency of the system. Nevertheless, the  $E_{rr}$  comprises both the interferences coming from the other zone as well as any modification of the reproduction characteristics (this is re-addressed later).

In addition, the target pressure definition is also used to define the reference acoustic contrast ( $RAC$ ) as the  $AC$  obtained when only  $P_R$  and  $P_L$  are driven in phase; *i.e.* the  $AC$  generated thanks to the distance of these sources to the zones (close to the  $BZ$  and far from the  $DZ$ ) and their interferences.

Lastly, in this chapter the  $RAC$  is used to calculate the acoustic contrast gain ( $AC_G$ ) as

$$AC_G(\omega) = AC(\omega) - RAC(\omega). \quad (3.1)$$

Consequently, the  $AC_G$  in this chapter quantifies whether if the system is improving the  $AC$  or not, *i.e.* evaluates the “acoustic separation” earned by the designed system.

## 3.2 Free-field conditions

The first step of this study is to assume ideal free-field conditions as environment. In spite of being conditions far from a car cabin, it provides an overview on how the system interacts with itself ignoring the contribution of the environment. Therefore, the free-field conditions are studied both with simulations (Complexity №1) as well as laboratory measurements (Complexity №2).

### 3.2.1 Complexity №1: Monopoles in free-field

The simplest level of complexity (№1) to assist the design of the system is to use theoretical monopoles in free-field. The IRs are calculated as integrer delays using

$$h_{k,l}[n] = \frac{\delta(n - \frac{\Delta_{k,l}}{c})}{\Delta_{k,l}}, \quad (3.2)$$

where  $h_{k,l}$  is the IR between microphone  $k$  and the source  $l$ ,  $\Delta_{k,l}$  the distance between microphone  $k$  and the source  $l$ ,  $\delta[n]$  is the unit sample and  $c_0$  the speed of sound.

All the IRs are calculated using a sampling frequency  $fs = 44100$  Hz and 0.2 s long IRs and, as a frequency domain formulation is being used, the Fourier transform is applied to all the IRs. To study the frequency band of interest, the optimisation is calculated between 0 Hz and 1500 Hz with a 5 Hz resolution.

#### 3.2.1.1 Results: the starting point

Firstly, one proceeded to calculate the *RAC* using monopoles in free-field (Fig. 3.2).

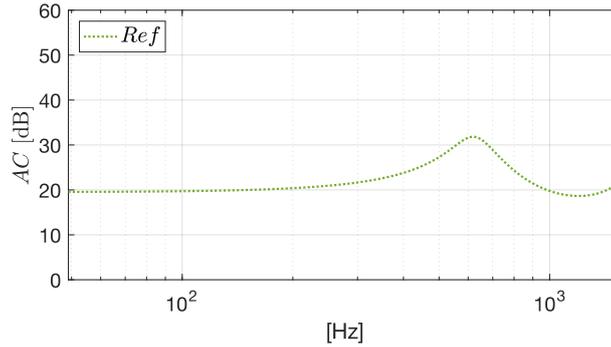


Figure 3.2: Resulting *RAC*, Monopoles in free-field.

Even if no treatment is applied ( $P_R$  and  $P_L$  in Fig. 3.1 driven in phase), a minimum *AC* of  $\simeq 20$  dB is achieved for the complete bandwidth (see Fig. 3.2). In spite of its redundancy, it is important to highlight that the levels of  $E_{rr}$  and an  $E_{ff}$  for the reference are always equal to zero because the target pressure is by definition the reference’s resulting pressures.

It is possible to notice that the *RAC* follows a “resonant-like” behaviour, more precisely, if a larger bandwidth was studied it is a “comb-like” behaviour. The *RAC*’s shape can be determined by looking at the target pressure’s resulting pressure in the *DZ*. For example, it is possible to explain the local maxima and minima of the *RAC* by analysing the relationship between the wavelength ( $\lambda$ ) and the distance between the sources defined for the target pressure ( $\Delta_{P_R, P_L} = 28$  cm) as in

$$\frac{n}{2}\lambda = \Delta_{P_R, P_L}. \quad (3.3)$$

As both sources are being driven in phase, when  $n$  is odd they superpose destructively, *i.e.* there is an inherent attenuation of the pressure in the *DZ*, consequently, an *AC* enhancement. Conversely, when  $n$  is even, both sources superpose constructively, therefore there is an amplification of the pressure in the *DZ*; as a consequence an *AC* reduction. The *Ref*’s “shape” is from this point forward referred to as the “natural response”.

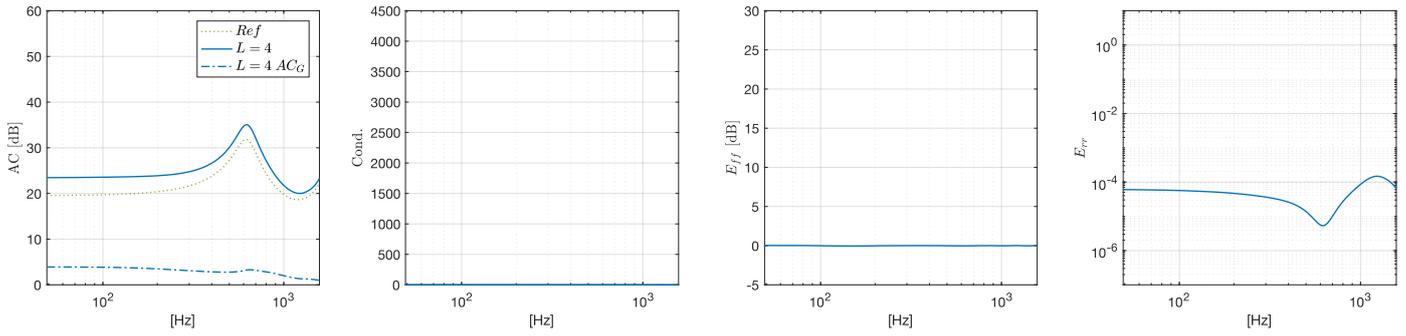


Figure 3.3: Resulting metrics employing four monopole sources in free-field.

When the PM algorithm developed in Sec. 2.3.1 is applied (blue curves in Fig. 3.3), it results in an  $AC_G$  of  $\approx 3.9$  dB for the lowest part of the spectrum, holding relatively constant up to 655 Hz and then decreasing with the frequency and becoming less than 1 dB above 1300 Hz. Regarding the other metrics, the  $E_{rr}$  results in values smaller than  $1.4 \times 10^{-4}$ , the effort is close to zero and the problem is far from ill-conditioned ( $\text{Cond.} < 2$ ) for the complete frequency band.

It is possible to observe in Fig. 3.3 that the natural response of the system is defining both the shape of the  $AC$  as well as the shape of the  $E_{rr}$ . In other words, the natural response of the system, imposed by the desired pressure, is strongly conditioning the performance of the system.

Additionally, to complete the study of the  $L = 4$  system the frequency response of the resulting filters are analysed, both amplitude and phase.

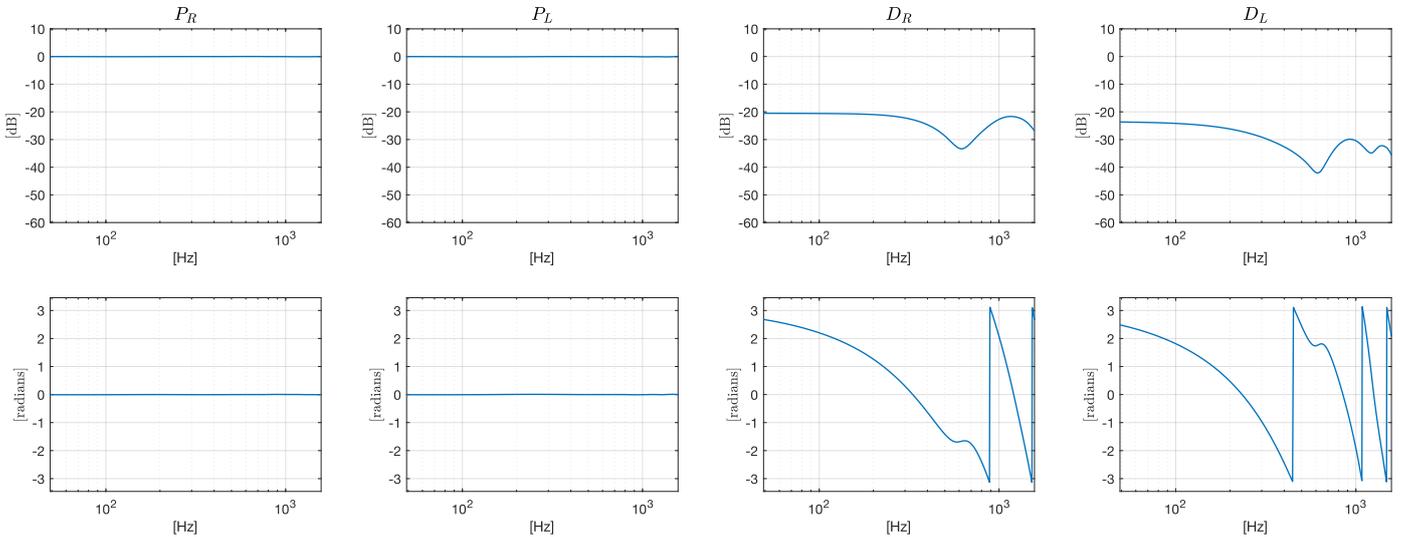


Figure 3.4: Resulting filters, four monopole sources in free-field.

As it was observed in [125], due to the proximity of  $P_R$  and  $P_L$  to the  $BZ$  and the fact that the target pressure is defined as the pressure when these sources are driven in phase, the filters associated to  $P_R$  and  $P_L$  are a copy of the ones used to define the target pressure, almost perfect all-pass filters with zero delay, namely, the frequency response of a  $\delta[n]$  function (see Fig. 3.4). Hence, it is possible to state that  $P_R$  and  $P_L$  are the largest -if not the sole- contributors to the “reproduction” part of the problem, *i.e.* the  $BZ$ .

Consequently,  $D_R$  and  $D_L$ , much closer to the  $DZ$  and far away from the  $BZ$ , run the “cancellation” part of the problem. Both filters are in charge of introducing energy to the system in order to destructively interfere what is being generated by the reproduction ones. However, it is important to remember that

the generation of PSZ is a twofold problem, the cancellation sources will introduce energy to the system as long as it does not affect much the reproduction in the  $BZ$ .

Furthermore, it is possible to observe the imposition of the natural response of the system in the shape of the cancellation filters.

In other loudspeaker configurations, for example sources in a circle around the zones [7, 115] or planar arrays [97, 99], as all the sources contribute to both the cancellation in the  $DZ$  as well as the reproduction in the  $BZ$ , it is not possible to use this reproduction/cancellation analysis.

### Summary of the $L = 4$ system, Complexity №1

In spite of improving the  $AC$  with very low  $E_{rr}$  levels, very small effort and excellent conditioning, the  $L = 4$  system is not capable of achieving the wanted levels of  $AC$ .

The use of near-field sources allows to separate the sources into two categories: the “reproduction” sources which are the largest contributors in the  $BZ$  and the “cancellation” sources in charge of cancelling what’s arriving from the  $BZ$  to the  $DZ$ .

In addition, it has been shown that the output of the algorithm strongly depends on how the target pressure is seen from the  $DZ$ , *i.e.*, the algorithm is reacting to the relationship between the source location and the field constraints specified in the target pressure, which are directly represented by the  $Ref$  or as referred in this thesis: the natural response of the system.

In order to enhance the  $AC$  (from Fig. 2.13) there are three possibilities:

- Change the field constraint in such a way that the natural response already achieves higher  $AC$  levels. However, to do this it is necessary to modify the target pressure and the system will no longer fulfil the transparency constraint.
- Change the optimisation *e.g.* add a weighting factor to change the balance between the potential energy and the error minimisation or directly using a different cost function, but as reviewed in Sec. 2.3.3.4, this deteriorates the  $E_{rr}$ , compromising the transparency of the system.
- Modify the inputs: *e.g.* adding more sources.

### 3.2.1.2 Adding sources: from $L = 4$ to $L = 6$

As it was mentioned when introducing the CO methods (in Sec. 2.2.3), the optimisation of the cost-function comprises both the numerical optimisation and the input optimisation. Hence, starting from the  $L = 4$  system, the addition of an extra source (an  $L = 5$  system) is studied in order to improve the minimisation of the cost function.

Therefore, to find the best location for the fifth source, a grid of 336 possible-positions was built defining a rectangle of  $1.37 \text{ m} \times 0.65 \text{ m}$  with a separation of 5 cm between each other (candidates in Fig. 3.5). After evaluating several heights, one concluded that the most relevant one -with the largest effect- was at the same height as the microphones and sources.

To quantify the contribution of the fifth source, both terms of the chosen cost function (Sec.2.3.3) are plotted as a function of the position of the fifth source (Fig. 3.5).

- The difference in  $E_{rr}$  between the  $L = 5$  and the  $L = 4$  configurations.
- The ratio between the resulting potential energy in the  $DZ$  with the  $L = 5$  and the  $L = 4$  configurations.

In spite of the fact that the PSZ system’s performance is frequency dependent (as shown in Fig. 3.3), from the  $L = 4$  configuration a broad-band improvement is needed, hence, the mean value is used to observe the contribution of the extra source.

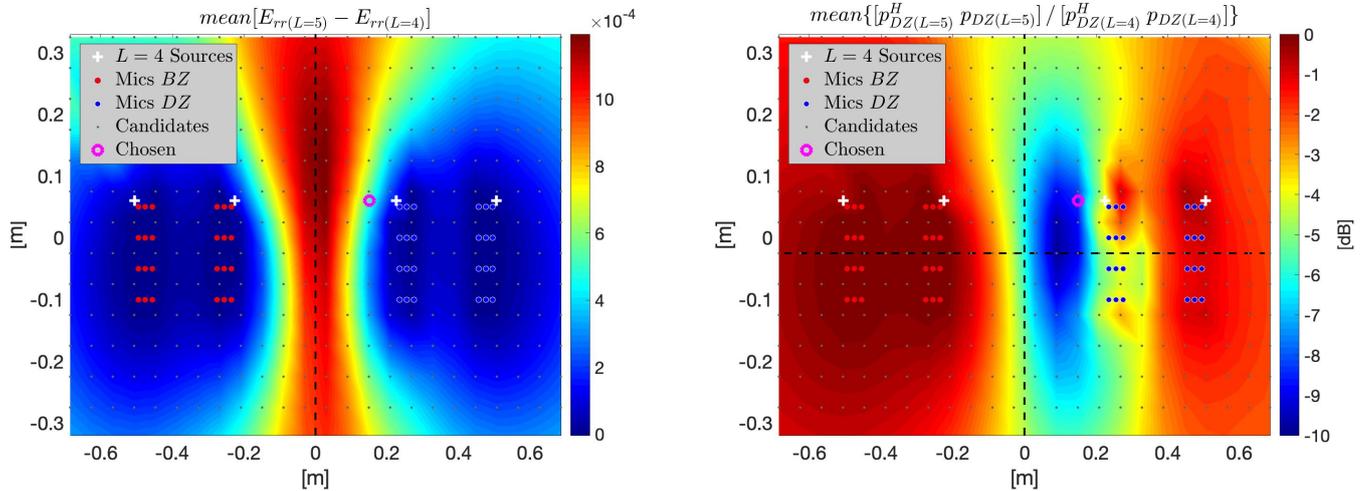


Figure 3.5: Study of the addition of a fifth source in free-field. Mean  $E_{rr}$  deterioration in the  $BZ$  (left) and mean energy reduction in the  $DZ$  (right).

Firstly, from Fig. 3.5 it is possible to state that the location of an extra source has a strong influence on the system’s performance.

On the one hand, the  $E_{rr}$  (Fig 3.5 left) presents a sort of vertical symmetry, where it is mainly deteriorated along the vertical centre line and decreasing outwards. The reason of this “sort-of-symmetry” is twofold: the further the source is from the  $BZ$ , the smaller the reproduction deterioration. Secondly, the closer the source is to the  $BZ$ , it is as if the algorithm “turns off” the fifth source because it strongly deteriorates the  $E_{rr}$  without contributing to the attenuation. Hence, the largest deterioration is in the middle line, where the deterioration of the  $E_{rr}$  is comparable to the potential energy reduction.

On the other hand, regarding the potential energy minimisation on the  $DZ$  (Fig. 3.5 right) the system no longer follows the vertical symmetry, the largest attenuation happens between the vertical centre line and the  $DZ$ . Nevertheless, it is possible to observe that between the zones a sort-of-horizontal “symmetry”, defined by the centre of the zones where the attenuation gets its maxima and decreases the further the source is from it.

Although the best minimisation is achieved when the fifth source is placed in this horizontal-line, it is necessary to have in mind the final application; *i.e.* it is not realistic to install a fifth source where it might obstruct the visibility of the driver. Therefore, the fifth source positioning (indicated with a magenta ring in Fig. 3.5) is chosen as a trade-off between performance and application. Firstly, to be on the same line as the original headrest-sources not to affect the visibility of the driver, secondly, from the centre 15 cm towards the  $DZ$  to reduce the  $E_{rr}$  deterioration but far enough from the  $DZ$  to have a larger impact on the attenuation. This position provides a mean attenuation of  $\approx -8$  dB with a deterioration of  $\approx 5 \times 10^{-4}$ .

One cannot help but compare these results with the ones published by Elliott and Jones in 2006 [3] where they also reviewed the addition of an extra source in a headrest configuration. But two main differences set these two analyses apart: on the one hand, the aim of Elliott’s paper was to actively generate a directive source so it is only considering the contribution of two sources and not the complete array, on the other hand, the algorithms used in this paper do not consider the reproduction error. Hence, the conclusions from this paper are not applicable in this study.

Lastly, to maintain a symmetric performance in both seats, the addition of a fifth source implies the addition of a mirrored sixth source as in Fig. 3.6, from this point forward known as the “inner-sources”.

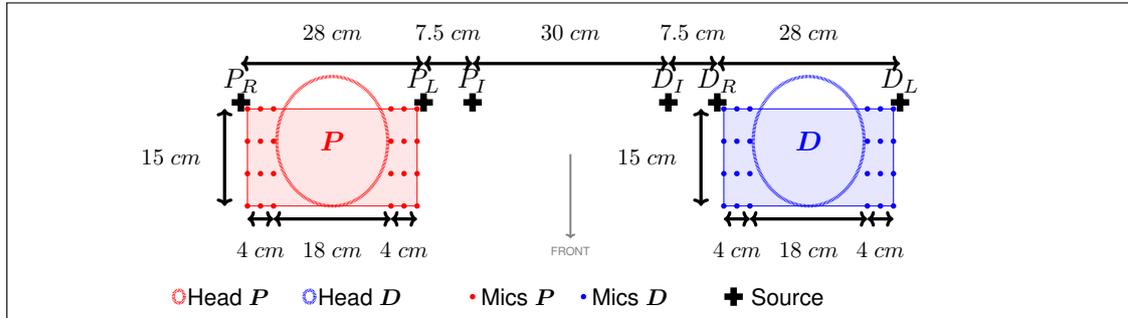


Figure 3.6: Schematic: six-source PSZ system.

Nevertheless, in order to evaluate the contribution of each inner source, the four metrics are analysed: as  $L = 5$  only adding  $D_I$  and as  $L = 6$  adding both inner sources  $P_I$  and  $D_I$ .

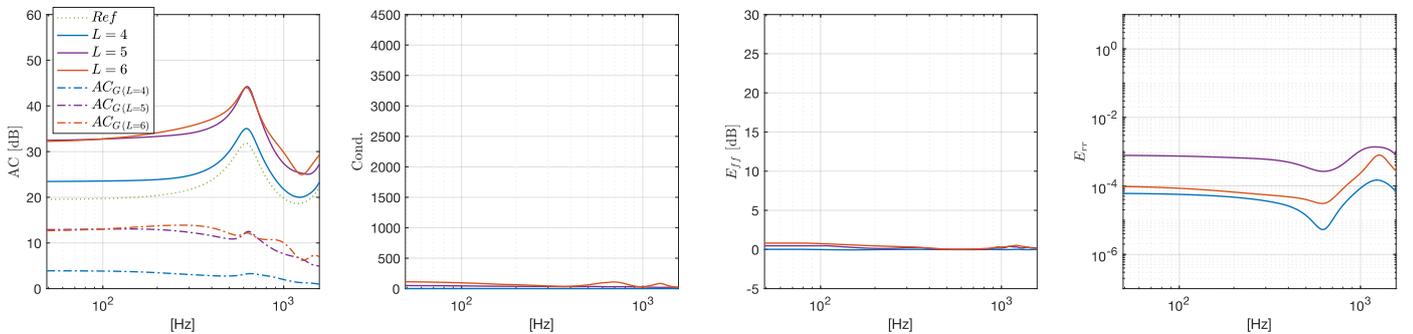


Figure 3.7: Resulting metrics employing four, five and six monopole sources in free-field. Performance contribution of the inner sources.

When only  $D_I$  is added ( $L = 5$ ), there is a considerable improvement in the  $AC$  performance, but there is a stronger deterioration of the  $E_{rr}$  (more than  $3.7 \times 10^{-4}$ ). However, when both  $P_I$  and  $D_I$  are used ( $L = 6$ ) the  $AC$ , one achieves almost the same level of  $AC$  but the deterioration of the  $E_{rr}$  is reduced.

Due to the fact that the target pressure definition has not changed, the resonant-like behaviour reviewed in Sec. 3.2.1.1 is still present in the  $L = 5$  and  $L = 6$  systems achieving levels of up to 47 dB of  $AC$  (around 612 Hz). Nevertheless, it is useful to observe the  $AC_G$  that stays above 10 dB for the frequency band of interest.

Lastly, one proceeded to analyse the  $L = 6$  resulting filters.

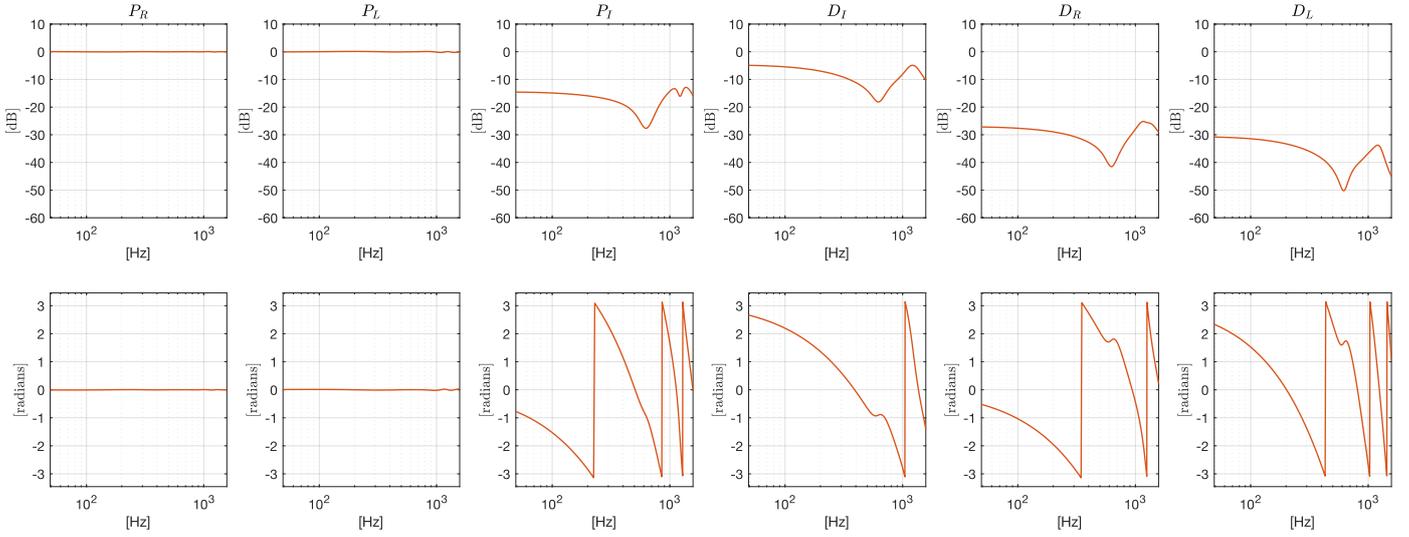


Figure 3.8: Resulting filters, six monopole sources in free-field.

Firstly, one can still easily identify the reproduction sources (in Fig. 3.5) in the  $L = 6$  configuration; both  $P_R$  and  $P_L$  are still almost perfect  $\delta[n]$  functions, therefore all the already mentioned advantages in  $L = 4$  can be extrapolated to  $L = 6$ . Analogously, it is possible to assume that  $D_R$  and  $D_L$  are solely contributing to the cancellation of the problem; they are both strongly attenuated and much closer to the  $DZ$  than any other source.

In spite of considering  $P_I$  and  $D_I$  also as cancellation sources ( $P_R$  and  $P_L$  are unquestionably the largest contributors to the reproduction), the role of  $P_I$  and  $D_I$  is twofold. On the one hand,  $D_I$  is resulting in the strongest contributor to improving the  $AC_G$  (3.7) whilst  $P_I$  it seems to introduce energy to the system in order to avoid  $D_I$  to deteriorate the reproduction. Nevertheless, this is just a hypothesis, unless the relative distances (sources/zones) is very different (as is the case of  $P_R$ ,  $P_L$ ,  $D_R$  and  $D_L$ ) their role in the overall problem becomes difficult to analyse and it might not be the same for every frequency.

### 3.2.1.3 Adding sources: from $L = 6$ to $L = 8$

In spite of fulfilling the desired  $\simeq 30$  dB, two extra “outer-sources ” are incorporated to the study (see Fig. 3.9). This larger configuration is chosen anticipating the presence of reflective surfaces -like a window or door- where it might be necessary to add more sources to control the field.

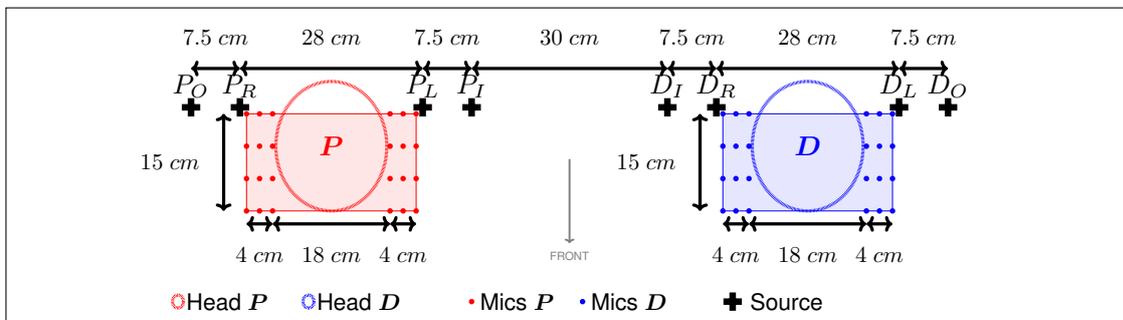


Figure 3.9: Schematic: eight-source PSZ system.

The performance of the system is calculated for the new system with the added outer sources ( $L = 8$ ).

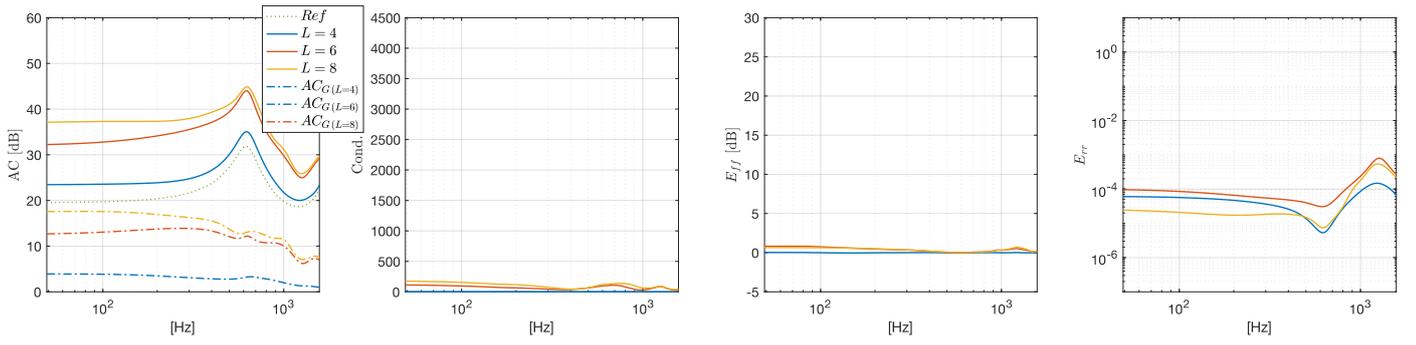


Figure 3.10: Resulting metrics employing four, six and eight monopole sources in free-field. Performance contribution of the outer sources.

In spite of not being as meaningful as the jump from  $L = 4$  to  $L = 6$  (Sec. 3.2.1.2), the addition of the outer sources also presents an improvement in performance (see Fig. 3.10). For the frequencies below the system’s “resonance” there is an  $AC$  gain as well as a reduction of the  $E_{rr}$  with a slightly larger condition number and an almost identical  $E_{ff}$ . However, above the system’s natural resonance, the contribution of the outer sources has almost no effect.

The filter’s frequency responses are also studied for the  $L = 8$  system (Fig. 3.11).

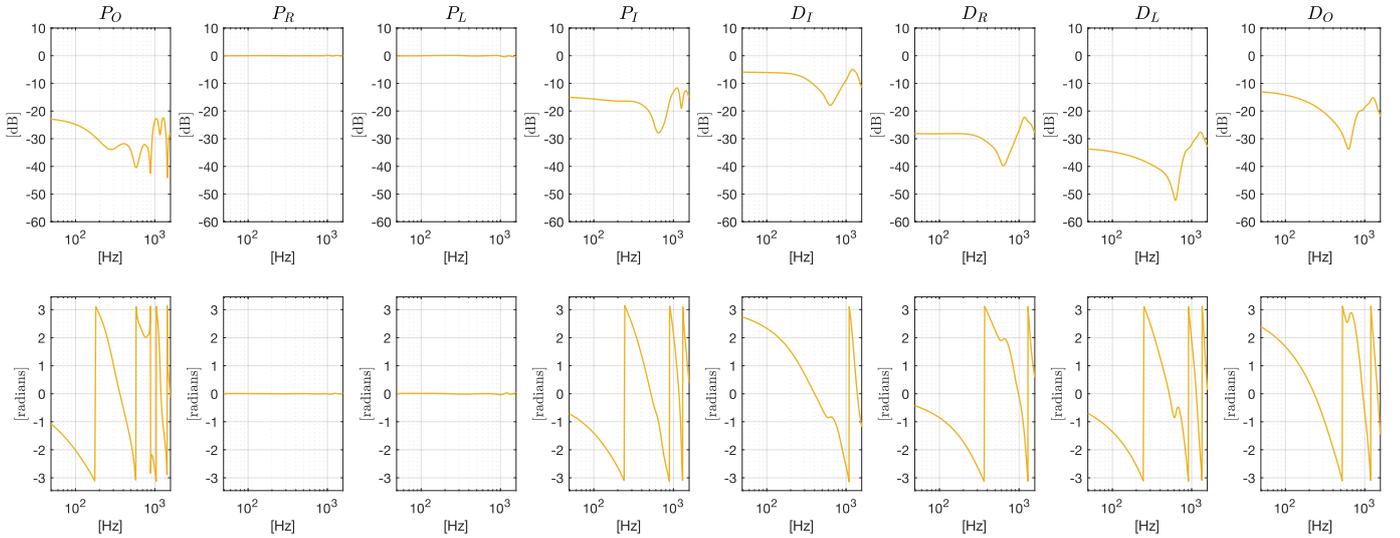


Figure 3.11: Resulting filters, eight monopole sources in free-field.

The analysis on the resulting for the  $L = 8$  configuration 3.11 is an extension of the ones made for the  $L = 4$  and  $L = 6$  configurations. Also in this configuration, the reproduction sources ( $P_R$  and  $P_L$ ) closer to the  $BZ$  correspond to the frequency response of the  $\delta[n]$  function. Moreover, the sources shared with the  $L = 6$  configuration share the same analysis, in spite of being more attenuated. The main difference is in the added  $D_O$  which is among the less-attenuated filters in the array.

When the outer sources are added, the algorithm is capable of obtaining a new filter combinations that improves the performance. Nevertheless, it is up to the manufacturer to choose if this improvement is worth spending the extra cost of two more loudspeakers, two more channels, associated cables, etc.

### 3.2.2 Complexity №2: Prototype in anechoic chamber

To validate the results obtained with the simulated data, as well as to increase the level of complexity, a prototype following the dimensions established in Fig. 3.9 was built. The system consists of two car-seats, separated by 73 cm with a dummy on each seat in order to emulate the presence of the passengers in a car (see Fig. 3.12).



Figure 3.12: Prototype Anechoic chamber.

One of the advantages of the system designed in the simulations (Fig. 3.9) is that it allows to build a compact system that contains all the sources in the headrest. Therefore each seat is provisioned with a custom headrest equipped with eight closed-box loudspeakers using the Tectonic TEBM35C10-4 LS (see Appendix C) with an enclosure of  $\approx 1$  litre each (see Fig. 3.13 left). The sources are all placed following the dimensions of the  $L = 8$  system provided in Fig. 3.9.



Figure 3.13: Prototype: Loudspeaker enclosure top view (left), seat close up (right).

The schematics used for the construction of the prototype with the all the dimensions can be found in

## Appendix B.

Lastly, the zones were sampled using 16 Brüel & Kjaer 4958 microphones arranged in groups of four per ear (see Fig. 3.13 right) and moved horizontally every 2 cm in order to obtain the same positioning as in Fig. 3.6. All the IRs were measured using a 17th order MLS algorithm averaged 24 times using a sampling frequency  $f_s = 44100$  Hz.

## 3.2.2.1 Complexity №2: Results

The first degree of acoustic complexity studied in this thesis allows to have a first overview on the limitations of the physical system and to design a system capable (at least under those conditions) to attain the original requirements. However, one of the assumptions that limits these simulations from the real application is the effect of having scattering surfaces inside and in the vicinity of the zones -*e.g.* a listener, the headrest, the seats, etc.- an issue already established as non-trivial in the literature [5, 87]. Therefore, as second degree of acoustic complexity, the prototype introduced in Fig. 3.12 was installed in the LAUM's anechoic chamber in Le Mans, France.

Using this system, all the IRs were acquired and the PSZ algorithm was used to calculate the *Ref*,  $L = 4$ ,  $L = 6$  and  $L = 8$  systems (Fig. 3.14).

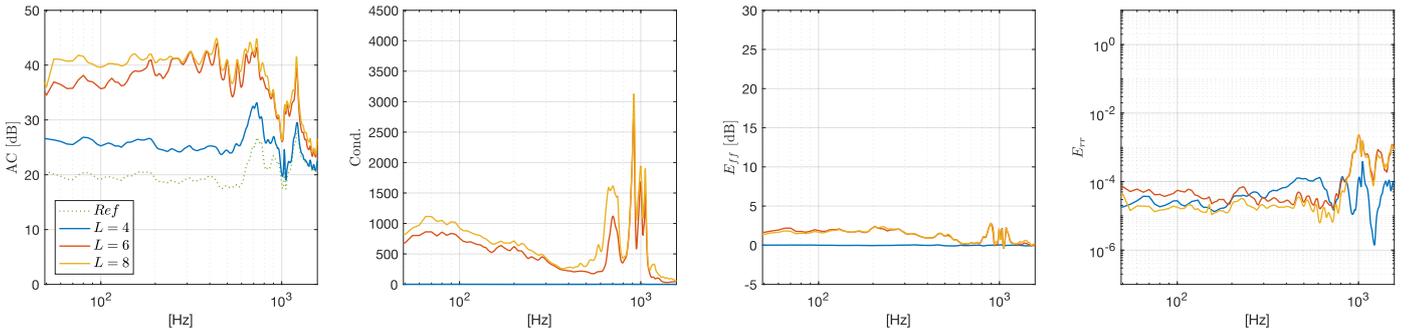


Figure 3.14: Resulting metrics employing four, six and eight loudspeakers in the anechoic chamber.

The first conclusion from Fig. 3.14 is that already the  $L = 6$  system is capable of achieving the 30 dB of *AC* whilst maintaining a small  $E_{rr}$  and  $E_{ff}$  (the  $L = 8$  only shows some improvement below 200 Hz).

When analysing the *RAC* in Fig. 3.14, it is possible to observe a similar resonant-like (comb-like) behaviour as in the monopole simulations (Fig. 3.7). The overall *RAC* is lower because the resonance presents a more “selective” effect. In spite of maintaining the same dimensions as in the simulations, the *RAC* “peak” is shifted to  $\simeq 720$  Hz, one of the reasons of this might be due to the difference between the position of a monopole and the real acoustic centre of a baffled loudspeaker [141] which could be measured and added to the monopole simulations, but this analysis is out of the scope of this thesis.

Regarding the *AC*, it is possible to draw the same conclusions as in the simulated conditions, the  $L = 4$  system is not capable of fulfilling the required 30 dB whilst the  $L = 6$  system goes beyond this threshold. It is possible to observe a drop in the *AC* at 1 KHz (in all three studied cases). This loss in performance is not due to the comb-like behaviour resulting from the target pressure definition, otherwise taking the *Ref* resonance value and following eq. 3.3, the fall should be at  $\approx 1400$  Hz. At this frequency,  $\lambda$  is comparable to the dummy, the headrest, etc. However, due to the complexity of the system ( $L = 6$  and 48 mics) it is not simple to explain this loss in performance.

The error in both the  $L = 4$  and  $L = 6$  configurations result in comparable performances with levels of the order of  $10^{-4}$ . In addition, the effort presented very similar results to the simulated case, with an effort  $\leq 3$  dB. On the other hand, there is a considerable growth in the condition number, but it is still considered low and the calculated filters were not affected by these numbers.

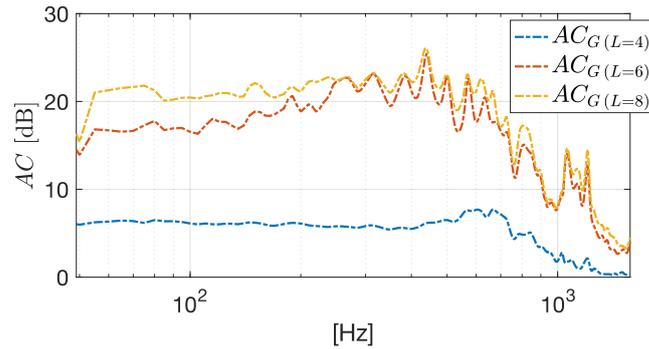


Figure 3.15: Resulting  $AC_G$ , employing four, six and eight loudspeakers in the anechoic chamber.

Moreover, when analysing the  $AC_G$ , the  $L = 4$  system is in the order of 5.5 dB up to the “resonance” where it starts falling consistently. On the other hand, the  $L = 6$  presents a  $AC_G$  of almost 20 dB up to  $\approx 715$  Hz where it also starts falling arriving to a value of 8.3 dB around 1 KHz. The underestimation of the  $AC_G$  in the monopole models (Sec. 3.2.1) is a direct consequence of the overestimation of the  $Ref$ .

It is possible to conclude that, in anechoic conditions, both  $L = 6$  and  $L = 8$  result in very similar performances but the second one represents a larger investment (two more sources, two more channels). Therefore, in anechoic conditions, the  $L = 6$  configuration is chosen.

Additionally to the studied metrics, one also analysed the resulting pressures over the zones for some specific levels of  $AC$  (see Fig. 3.16). In order to compare among the achieved attenuations, the resulting pressure is normalised for each system, for each frequency to the largest obtained pressure.

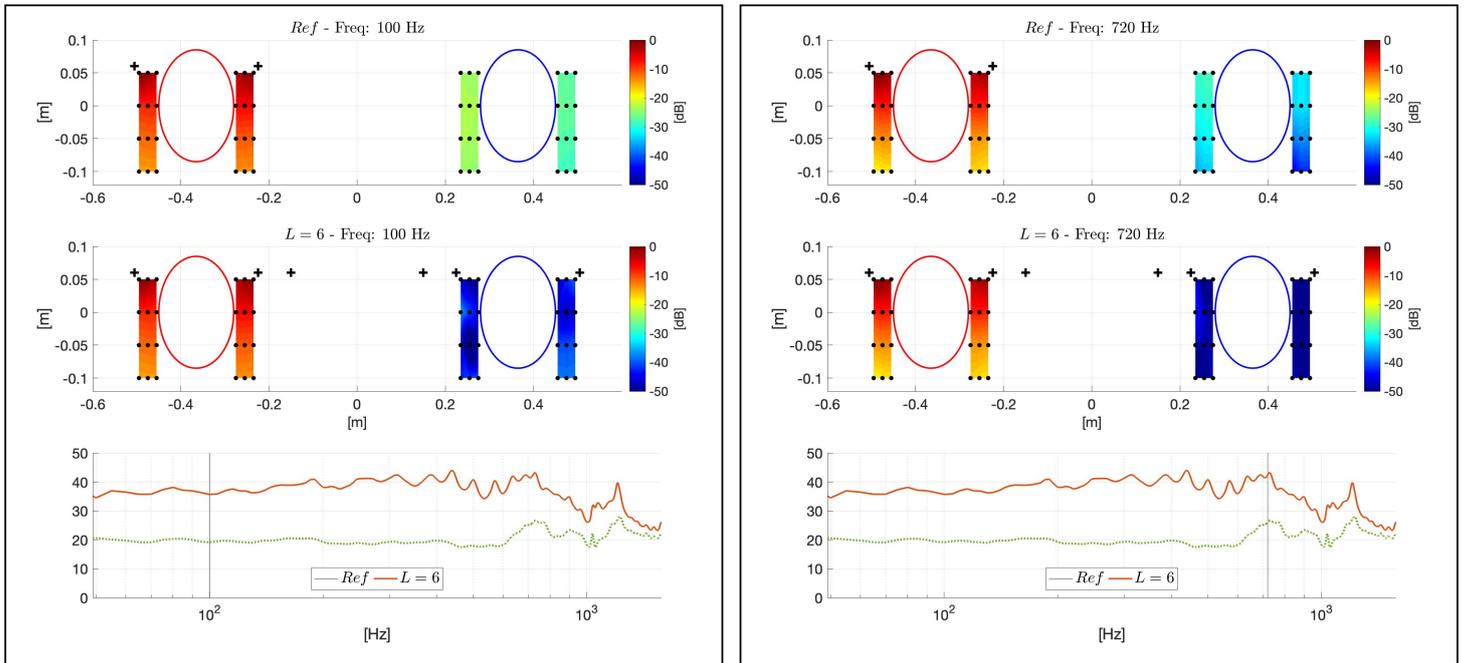


Figure 3.16: Normalised pressure for the  $L = 6$  configurations: At 100 Hz (left) and at 720 Hz (right).

Two main aspects can be seen in Fig. 3.16 when analysing the pressure over the  $BZ$ . Firstly, the modulus of the resulting pressure is the same for the  $Ref$  and for the  $L = 6$  system at both frequencies; just some rough evidence of transparency constraint. In addition, it is possible to observe how the pressure decreases as the microphone is further from the loudspeaker.

Regarding the pressure in the  $DZ$ , the  $L = 6$  at 100 Hz one can observe attenuations of the order of

40 dB and at 720 Hz attenuations of the order of 50 dB. One could say that the  $AC$  is underestimating the achieved-attenuation of the system, but it is important to remember that the  $AC$  is a metric that indicates the ratio of the potential energy between the zones and as mentioned the pressure over the  $BZ$  is not constant, it decreases with the distance. Lastly, when comparing the pressure in the  $DZ$  at both frequencies, one can observe that the  $Ref$  is considerably more attenuated for 720 Hz than it is for 100 Hz. This is due to the natural response of the system which -as previously mentioned- naturally attenuates the pressure over the  $DZ$ .

Lastly one proceeded to analyse the frequency responses of the resulting filters.

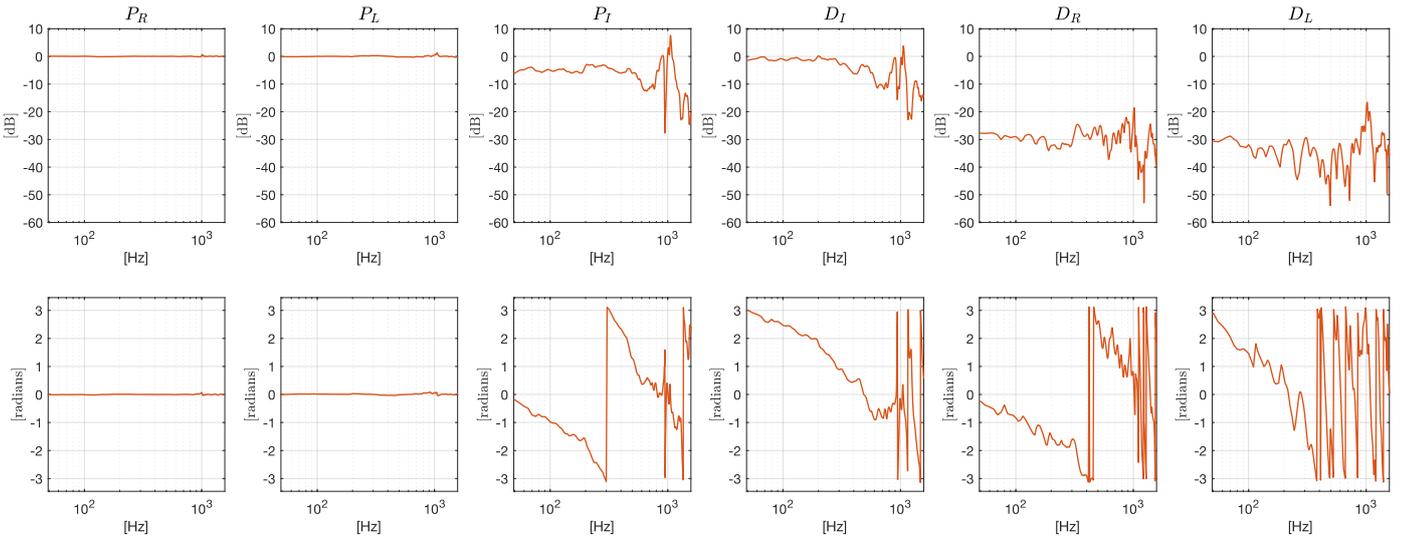


Figure 3.17: Resulting filters, six loudspeakers in the anechoic chamber.

Also in the measurements in the anechoic chamber,  $P_R$  and  $P_L$  are almost perfect  $\delta[n]$  functions. Furthermore, all the other sources are driven by more complex filters. The two sources closer to the  $DZ$  are the more attenuated and  $P_I$  and  $D_I$  are the sources contributing the most to the overall problem.

Below the so-called resonance at  $\approx 720$  Hz (taking the  $Ref$  in Fig. 3.14), it is possible to agree that the system presents a relatively smooth response. As the wavelength becomes comparable to the elements in the system (head, shoulders, headrest, etc.), the filter shapes become more complex with notch-type responses with more complex responses (*e.g.* at 950 Hz). Additionally, the notch-type responses coincide with the frequencies where there is a peak in the conditioning of the problem (see Fig. 3.14)

### Summary of Free-field conditions

Several conclusions are drawn from the first two levels of complexity: simulations and measurements in free-field.

Firstly, the use of near-field sources already puts the starting point of research at an  $AC$  of  $\simeq 20$  dB; *i.e.*, any system designed that does not achieve higher levels of  $AC$  than this are not worth the investment. In addition, by using the existing headrest sources and applying a PSZ one can already improve the performance by  $\simeq 5$  dB.

With the help of the monopole simulations, a headrest system was designed capable of achieving the wanted levels of  $AC$  with small  $E_{rr}$ . The ideal free-field with monopole sources was used to determine the source positioning and its performance was validated in laboratory conditions with real seats, with a prototype headrest and two dummies as listeners.

The way the target pressure is defined pre-conditions the algorithm (Field Constraint in Fig. 2.13) and its resulting  $AC$  is strongly affected by the natural response of the system.

The resulting performances of the system showed that the  $L = 6$  configuration already achieves the desired performance and that the addition of two outer sources ( $L = 8$ ) does not offer a big improvement, neither in the  $AC$  nor in the  $E_{rr}$ .

### 3.3 Controlled reflections

Once the desired performance is achieved under perfect-laboratory conditions, the environment's complexity is increased by adding a series of reflective surfaces. Matching the dimensions of a mid-size car, two parallel reflective surfaces each one at 68.5 cm from the centre of the system are introduced (as in Fig. 3.18).

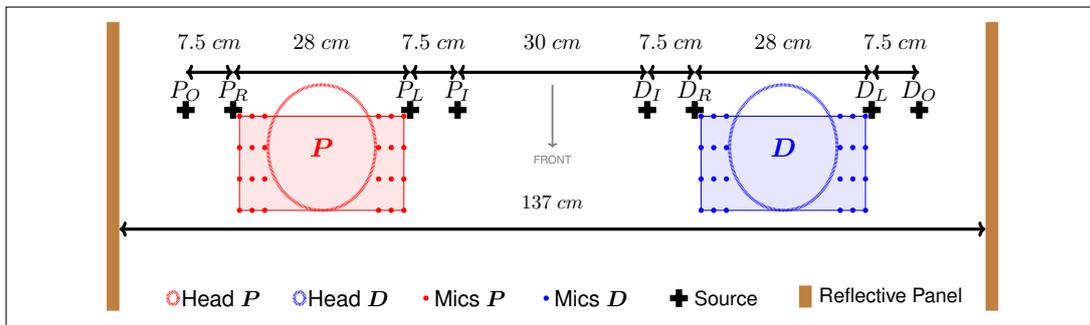


Figure 3.18: Schematic: eight-source PSZ system with reflective surfaces.

To evaluate the effect of the reflective surfaces on the algorithm and using the headrest-system designed in Sec. 3.2 three different environments are studied (See. Fig 3.18):

- 1) A reflective surface on the passenger side ( $P$ ), close to the  $BZ$ .
- 2) A reflective surface on the driver side ( $D$ ), close to the  $DZ$
- 3) Both surfaces simultaneously.

As done in free-field conditions (Sec. 3.2), firstly simulated IRs are used (Complexity N°3) and then validated experimentally (Complexity N°4).

#### 3.3.1 Complexity N°3: Monopoles with infinite-baffles

As third level of acoustic complexity, the sources are again considered to be perfect monopoles (IRs expressed with Eq. 3.2) but to take into consideration the reflective surfaces, a first order specular reflection is considered by placing a mirror source equidistantly from each reflective surface as if an infinite baffle was present (see Fig. 3.19).

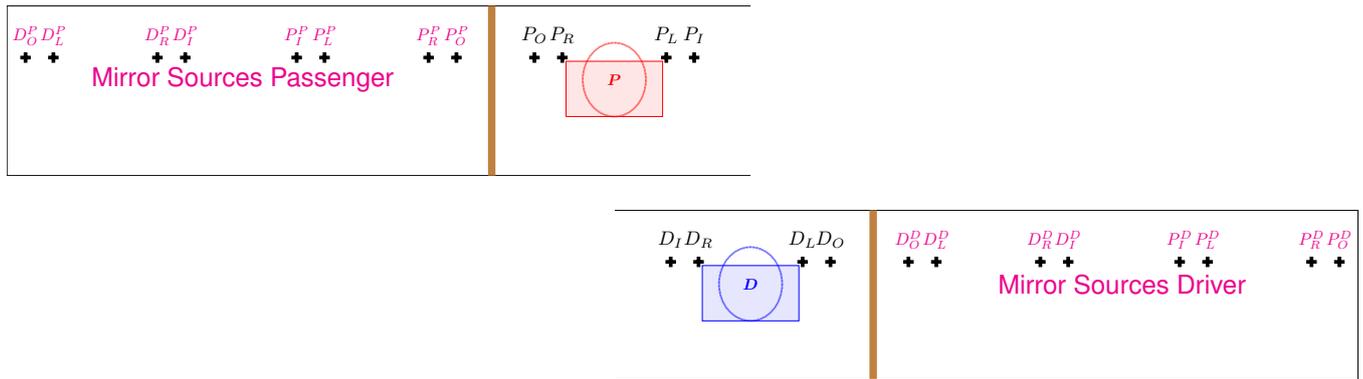


Figure 3.19: Reference system schematic with infinite baffles with first order mirror sources.

It is important to highlight that when both baffles are present, only a first order reflection is considered *i.e.* the presence of a standing wave is ignored in the simulations.

#### Notes on the reference contrast:

Before analysing the performance of the system, it is necessary to calculate the *Ref* (its natural response). As explained in Sec. 3.1, the *Ref* is defined as the *AC* naturally achieved by the system when only  $P_R$  and  $P_L$  are driven in phase. However, there is a not-so-obvious subtlety in this definition that needs to be addressed. It implies that different acoustic environments, in spite of using the same loudspeakers and zones, result in different *Refs*, consequently different natural response.

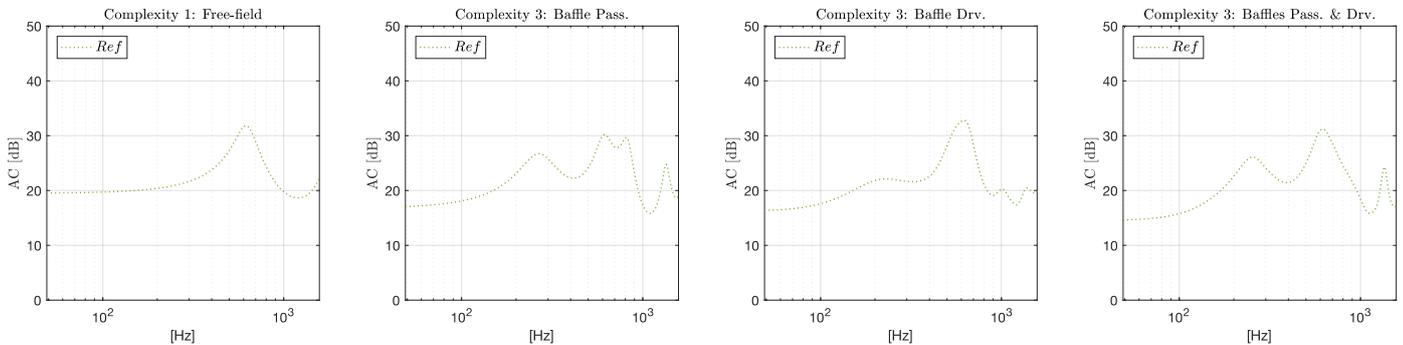


Figure 3.20: *RAC* in four different environments: Free-field, baffle passenger, baffle driver, passenger and driver baffles only considering first-order reflections.

Thus, when comparing the *Refs* in Complexity №2 with the one in Complexity №1, “new resonances” appear (see Fig. 3.20). Although these new peaks and depths are also a product of the resulting pressure field in the *DZ*, the *Ref* is no longer only ruled by the interferences between  $P_R$  and  $P_L$ , but also by all the interferences with all their mirror sources (see Fig. 3.19). Hence, by calculating all the distances between the sources and the distances between the sources and its mirror sources one could identify using Eq. 3.3 the frequencies of the peaks and depths in Fig. 3.20.

For example, if one wanted to anticipate the frequencies and whether the sources interfere constructively or destructively when the panel near the passenger is present, using Eq. 3.3 one could for example assemble a table as follows for the studied bandwidth

	$\Delta$	In Phase (loss $AC$ )	Out of Phase (enhancement $AC$ )
$P_R + P_L$	28 cm	1225 Hz	612 Hz
$P_R + P_R^P$	36 cm	952 Hz	476 Hz
$P_R + P_L^P$	64 cm	536 Hz, 1072 Hz	268 Hz, 804 Hz
$P_L + P_R^P$	64 cm	536 Hz, 1072 Hz	268 Hz, 804 Hz
$P_L + P_L^P$	92 cm	373 Hz, 746 Hz	186 Hz, 559 Hz, 932 Hz

Table 3.2: Distance between Monopoles and Mirror sources with baffle on passenger side.

From table 3.2, one can conclude that the more complex the environment the more complex the interaction between the sources (and its mirrors). Only studying the effect of  $P_R$  and  $P_L$  with one reflective panel we can identify 17 different interferences in the studied bandwidth. Moreover, Eq. 3.3 only allows to identify the frequency, but not every interference has the same amount of effect on the final  $Ref$ .

Lastly, it is important to remind that this change in the natural response from environment to environment (in this case depends on which reflective surface is considered), not only changes the reference to measure the  $AC_G$  (see Eq. 3.1) but also each environment constraints the PSZ algorithm differently.

### Baffle passenger side, near $BZ$

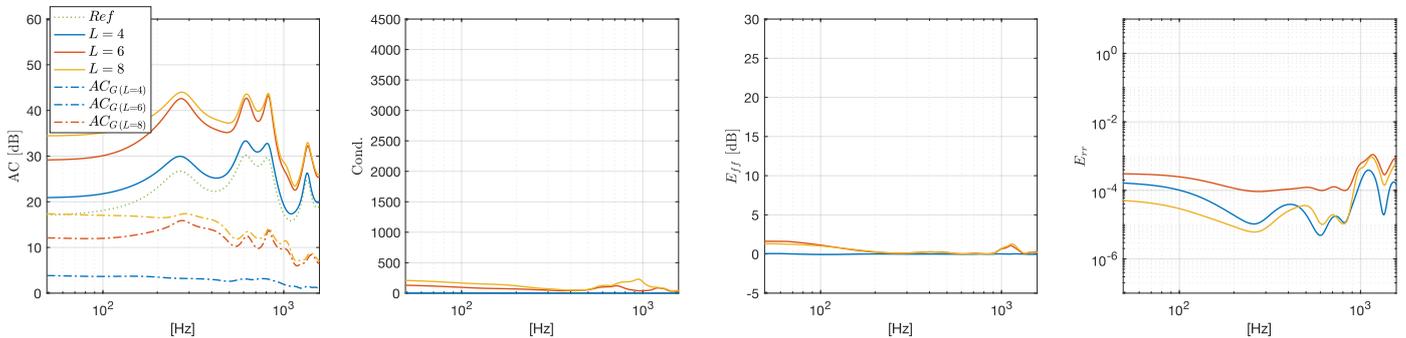


Figure 3.21: Resulting metrics employing four, six and eight monopole sources in free-field with an infinite baffle near the passenger.

When comparing the overall performance in presence of the baffle near the passenger (Fig. 3.21) with the results obtained in free-field conditions (Fig. 3.10), it is possible to find several common aspects. The four monopole configuration is improving the performance as in the anechoic case (no more than 3 or 4 dB) and it is following the same “shape” as the  $Ref$ . Furthermore, The addition of the inner monopoles proves to be as effective in the presence of the baffle and the  $E_{ff}$  and Cond. are very close to the anechoic case. Lastly, the  $E_{rr}$  when using  $L = 6$  sources is larger than the other two configurations (slightly more than seen in the free-field conditions).

Nevertheless, the main difference is the shape of the  $AC$  and the  $Ref$ . When the baffle close to the passenger is present, all the systems are capable of improving the performance but they are mimicking the response of the  $Ref$  with more (the  $L = 8$  system) or less (the  $L = 4$  system)  $AC$ . This is due to the fact that the reflective surface is so close to the  $BZ$ , the system is not capable of modifying its natural response, at least not without deteriorating the  $E_{rr}$ , even having a source ( $P_O$ ) placed between the baffle and the headrest.

### Baffle driver side, near $DZ$

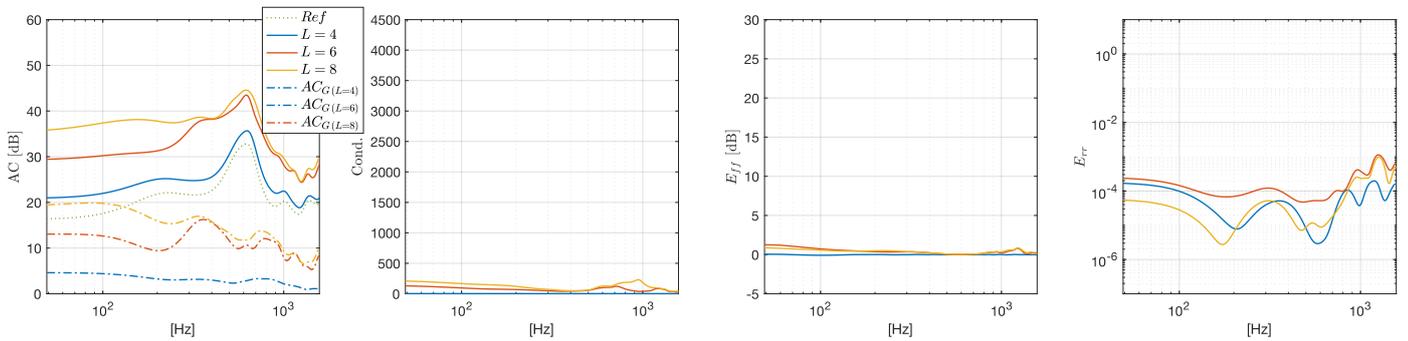


Figure 3.22: Resulting metrics employing four, six and eight monopole sources in free-field with an infinite baffle near the driver.

When adding only the baffle on the driver side, the  $L = 4$  configuration presents a 3-4 dB  $AC_G$  and follows the *Ref* almost identically. However, the other systems  $L = 6$  and even more the  $L = 8$  are showing the capability of “breaking” the natural response of the system whilst maintaining very low  $E_{ff}$  and Cond. Although the definition of the target pressure does not change, the fact that the reflective panel is far from the *BZ* not only decreases its interference in the *Ref* but also allows the system to introduce more energy to break its natural response without affecting much the  $E_{rr}$ .

In addition, there is a remarkable resemblance between the  $L = 8$ 's  $AC$  shape with the free-field conditions (Fig. 3.10); *i.e.*, in spite of having a reflective surface, the system is capable of minimizing the effects of these reflections (Fig. 3.21).

Nevertheless, this effect of having the reflective surface closer to the *DZ* also results in a strong modification on the  $L = 6$   $AC_G$  showing several peaks and depths, whilst the  $L = 8$  system presents a more robust  $AC_G$  especially below 300 Hz where it achieves free-field  $AC_G$  levels.

### Both baffles present

Lastly, the algorithm is evaluated when both baffles are present Right and Left (but only their first order reflections).

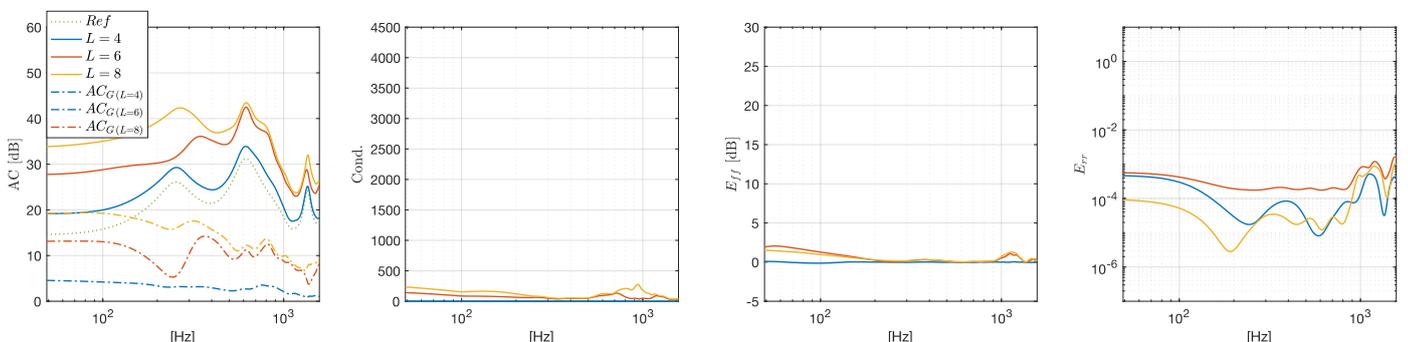


Figure 3.23: Resulting metrics employing four, six and eight monopole sources in free-field with two parallel infinite baffles (only first order reflection).

The overall performance remains analogous, the difference in the contribution of the  $L = 6$  compared to the  $L = 4$  is still present and it costs the increase of the reproduction error and the conditioning is held

very similar to the previous studied cases. Moreover, the  $L = 4$  system can not break the *Ref* shape and its gain remains in the order of  $\approx 4$  dB. On the other hand, the  $L = 6$  system is still capable of “breaking” the *AC* shape but its  $AC_G$  strongly depends on the interaction with the baffle closer to the *DZ*.

Lastly, the  $L = 8$  is chosen as the best compromise overall, it holds a more robust  $AC_G$  (more anechoic-like) and its capable of achieving levels of  $AC_G$  of more than 10 dB for the bandwidth of interest with smaller  $E_{rr}$  than the  $L = 6$  configuration (and in several frequencies even smaller than the  $L = 4$ ).

Then, frequency response of the resulting filters for the  $L = 8$  system is analysed.

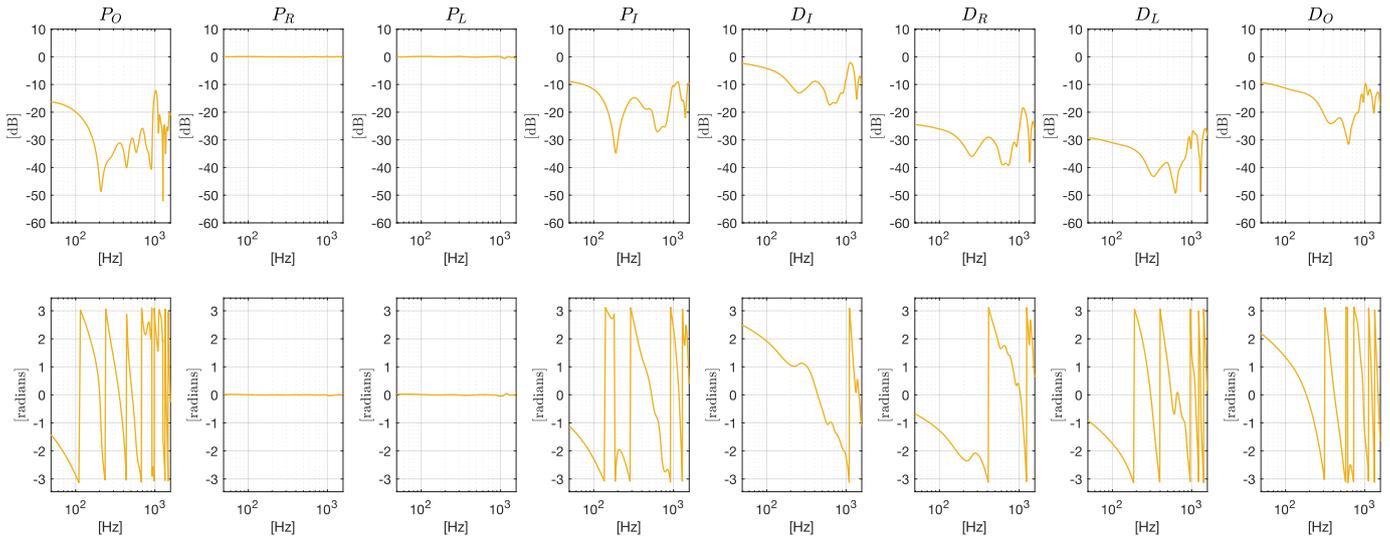


Figure 3.24: Resulting filters, eight monopole sources with two parallel baffles.

To analyse the resulting filters, it is possible to help oneself by the introduced reproduction/cancellation analysis. Once more, the reproduction sources  $P_R$  and  $P_L$  are easily identified, as they result again in the FRF of a  $\delta[n]$  function. On the other hand, among the cancellation filters it is possible to distinguish several different behaviours, nevertheless, their function is the same, to control what is being generated by the reproduction sources without affecting the  $E_{rr}$ .

As in the anechoic conditions (Fig. 3.11, both sources closer to the *DZ* ( $D_R$  and  $D_L$ ) are strongly attenuated ( $\leq 30$  dB). Moreover, the loudest loudspeaker is  $D_I$  which was showed in anechoic conditions to be the largest booster of the *AC* followed by the  $P_I$  that was hypothesised to be the one keeping  $D_I$  from “polluting” the *BZ* (due to its improvement in the  $E_{rr}$  Fig. 3.5). Lastly, the two outer speakers present very dissimilar behaviours. On the one hand, it is possible to assume that  $P_O$  is not contributing much to the overall problem, because its strongly attenuated and far from the *DZ*. However,  $D_O$  is unquestionably contributing to the cancellation problem.

Overall, the cancellation filters are working on the same order of magnitude as in free-field but, as the environment becomes more complicated, their responses start becoming more complicated too; more jumpy phase responses and more and steeper notch-type depths. Nevertheless, the reproduction filters regardless of the environment stay constant as almost perfect FRF of a  $\delta[n]$  function.

### 3.3.2 Complexity N°4: Prototype with reflective panels

In order to validate the simulated results of Sec. 3.3.1 and increasing the level of complexity, one carried out a series of measurements using the prototype introduced in Sec. 3.2.2 with two  $207 \text{ cm} \times 140 \text{ cm} \times 2 \text{ cm}$  medium-density fibreboard panels as reflective surfaces following the dimension in Fig. 3.18.



Figure 3.25: Prototype in the anechoic chamber with reflective panels.

In addition to the study of the reflections, Complexity №4 also considers the diffraction from the dummy, the seat and headrest as well as the use of loudspeakers.

### Both driver and passenger panels

The conclusions drawn from the measurements are very similar to the ones obtained with the first-order monopole simulations (Sec. 3.3.1), for this reason one is going to only analyse the most complex condition, when both panels are present.

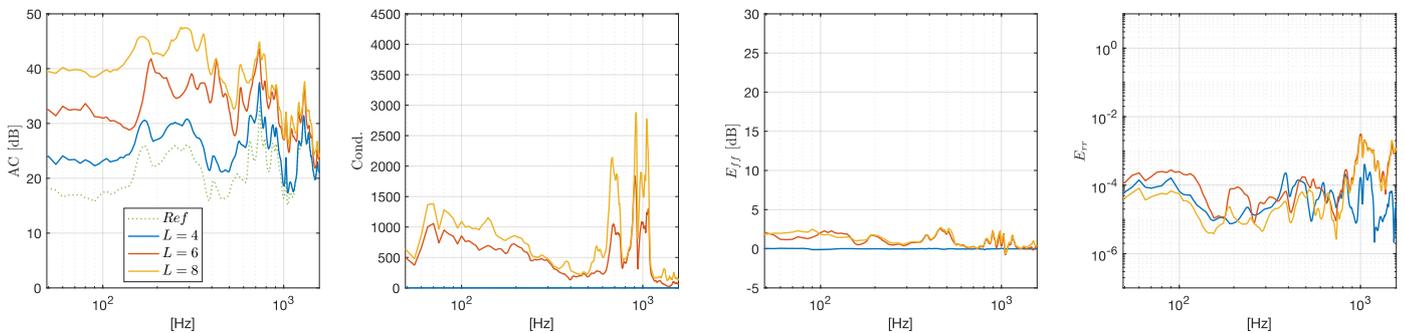


Figure 3.26: Resulting metrics employing four, six and eight loudspeakers in the anechoic chamber with two parallel panels.

Many of the conclusions obtained with the first-order reflections can be directly applied in the measurements. The  $L = 4$  system follows the natural response of the system, whilst the other systems are capable of breaking this shape and achieving larger levels of  $AC$  (especially  $L = 8$ ).

As it was analysed in Fig. 3.20, the *Ref*'s shape is the result of the peaks and depths caused by both  $P_L$ ,  $P_R$  and all their reflections. In order to identify each of the peaks and depths in the *Ref* it would be necessary to estimate the acoustic centre correction of the sources and add it when calculating the distances, not only between the sources but also their mirrors.

In addition, Complexity №3 assumes, not only perfectly rigid infinite baffles, but also only first order reflection ignoring the possibility of a standing wave. For example at  $\approx 540$  Hz, there is a drop in the  $AC$  for both the  $L = 6$  and  $L = 8$  which was not present in the original *Ref*. This frequency happens to coincide with the mode  $0 - 4 - 0$  which places both inner sources ( $P_I$  and the  $D_I$ ) in nodal lines reducing its capabilities to control the field.

Nevertheless, the general behaviour in the presence of controlled reflections is described in first-order simulations of Sec. 3.3.1. When using the  $L = 6$  system it is possible to observe an increase on the  $AC$ , but the addition of the outer sources ( $L = 8$ ) outperforms the other systems arriving to the aimed  $\geq 30$  dB. The main difference between the simulations and the measurements is regarding the  $E_{rr}$ , where all three systems result in comparable performances (at least below 900 Hz), *i.e.* the deterioration when using  $L = 6$  is smaller than simulated.

On the other hand, it is possible to observe a growth in the general condition number, especially around 1 kHz. As mentioned before, as the wavelength becomes comparable to the elements introduced in the system (dummy, headrest, etc.) the problem to solve becomes more complicated, hence the worst conditioned problem. Nevertheless, this levels of conditioning did not present any numerical problems.

From the resulting  $AC$  and the  $RAC$ , one calculated the  $AC_G$  for the four systems (Fig. 3.27)

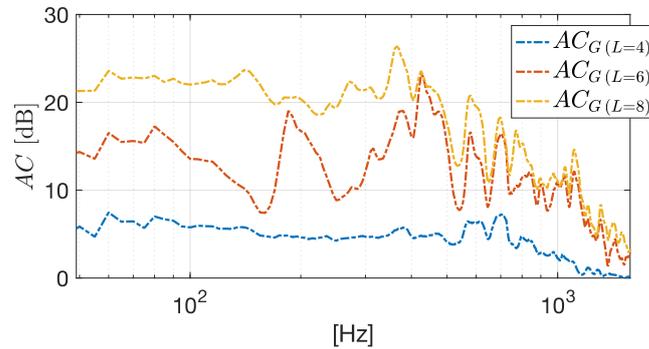


Figure 3.27: Resulting  $AC_G$ , employing four, six and eight loudspeakers in the anechoic chamber with two parallel reflective panels.

When observing the  $AC_G$ s in Fig. 3.27 is possible to draw the same conclusions as in Sec. 3.3.1, the addition of the outer sources result in a larger and more anechoic-like shape, whereas the  $L = 6$ 's  $AC_G$  strongly depends on the environment. In addition, the  $L = 8$  guarantees an  $AC_G \geq 10$ dB for the bandwidth of interest.

Therefore, one proceeded to study the resulting filters for the  $L = 8$  system when both panels are present (Fig. 3.28).

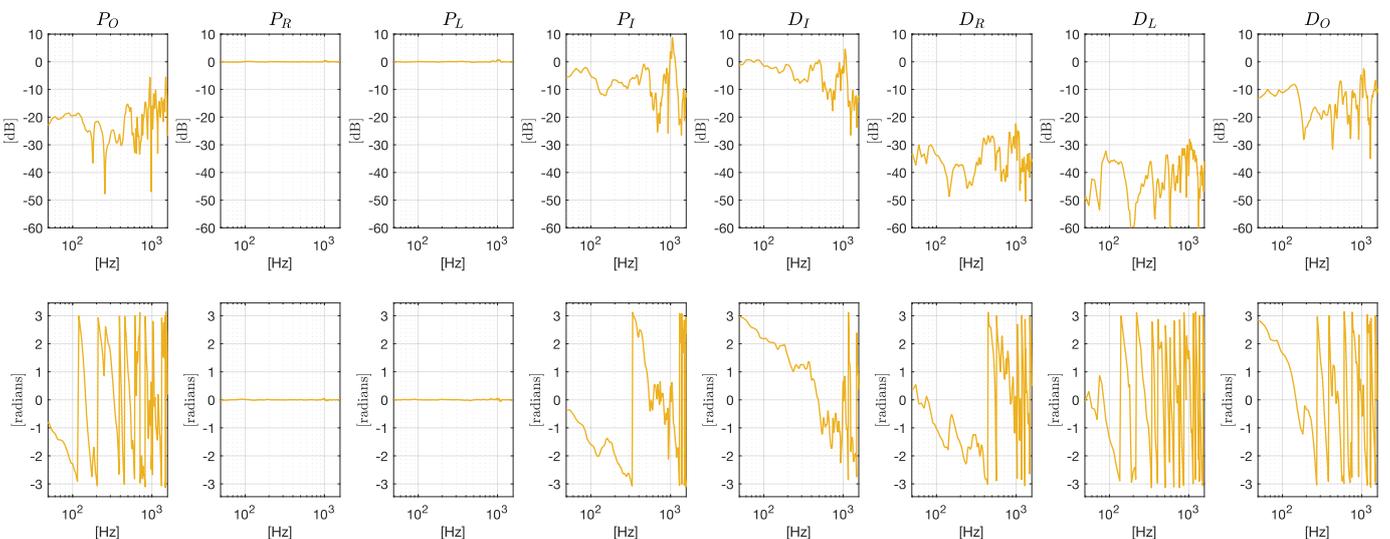


Figure 3.28: Resulting filters, eight loudspeakers in the anechoic chamber with two parallel panels.

Once more, it is possible to identify the reproduction sources  $P_R$  and  $P_L$  that result in almost perfect FRF of the  $\delta[n]$  function. In addition, it is possible to assume that most of the cancellation problem is being handle by  $P_I$ ,  $D_I$  and  $D_O$  and probably due to its proximity to the  $DZ$  a local contribution from  $D_R$  and  $D_L$ .

It is possible to make an analysis removing sources step by step and reviewing their punctual improvement to the overall performance similar to what was done in the source addition part (Sec. 3.2.1). However, it is important to remember that one of the reasons the PSZ were implemented is because the problem is too complicated to be solved analytically (optimise over zones and not points, frequency dependent, many sources, etc.). For example, at 430 Hz the contribution from  $D_O$  is very small (a gain of  $-31$  dB) so one could assume that it is dispensable, but at 540 Hz the system happens to coincide with the already mentioned mode  $0 - 4 - 0$ , where this source is providing an important contribution of the overall system with a gain of  $-6.7$  dB.

Therefore, the  $L = 8$  system is chosen as the best trade-off between cost (number of channels) and performance; it offers the algorithm a fair number of possibilities to fulfil the industrial constraints.

### Bonus: Passenger + Driver + Ceiling panels

The conditions analysed so far has kept the interactions between the environment and the system in 2D (in a plane). However, in a car cabin the contributions from the environment will come also from the ceiling and the floor. Therefore a third reflection panel was added to the already studied mock-up as shown in Fig. 3.29.

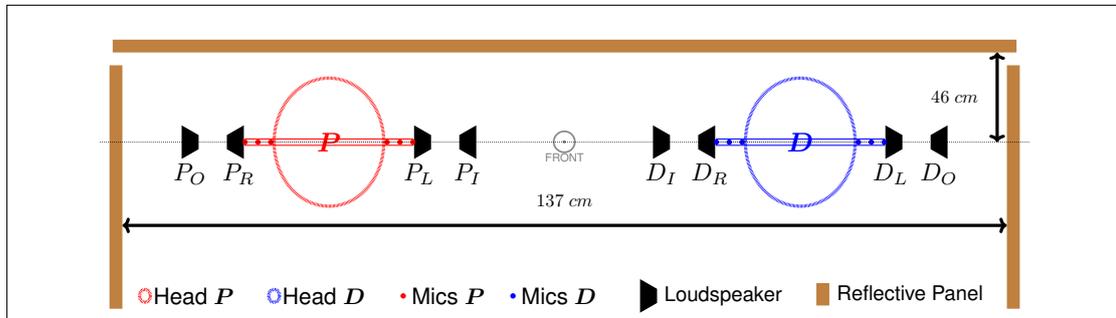


Figure 3.29: Eight-loudspeaker system front view with three panels, two parallels on the side and one on the ceiling. Measured configuration (above), general schematic with dimensions (below).

The third reflective panel in the ceiling is introduced in the anechoic chamber to the already studied prototype as in Fig. 3.29 and the metrics evaluated in Fig. 3.30.

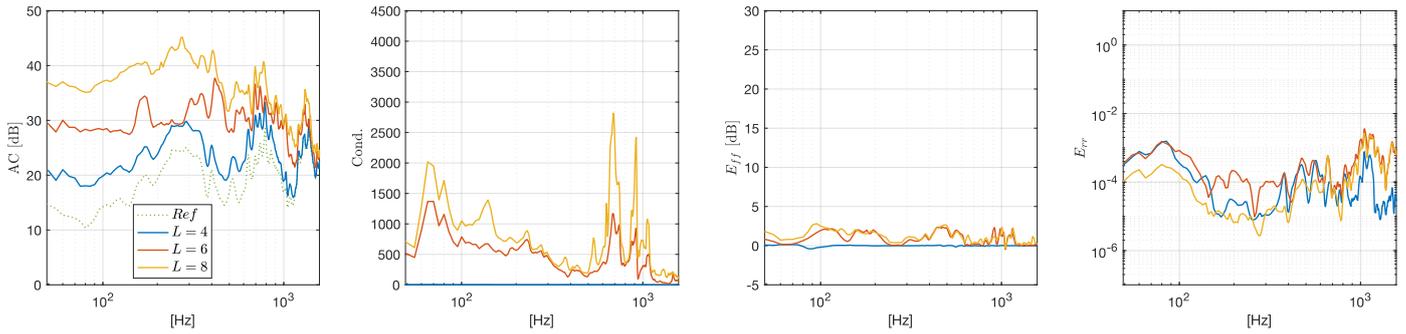


Figure 3.30: Resulting metrics employing four, six and eight loudspeakers in the anechoic chamber with two parallel panels and one in the ceiling.

Analysing the results obtained with the three panels (in Fig. 3.30), the same conclusions can be withdrawn. Especially the fact that the  $L = 8$  presents the most robust case overall achieving almost the same levels of performance as if the ceiling panel was not present (see Fig. 3.26).

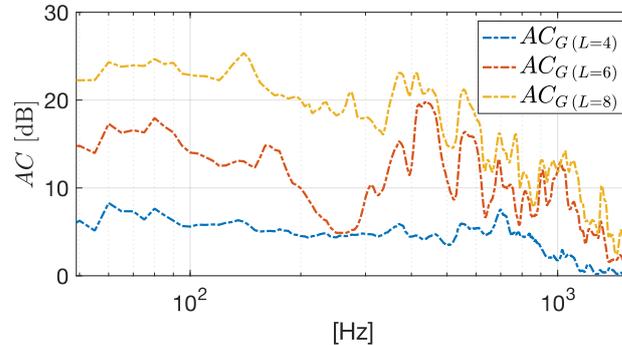


Figure 3.31: Resulting  $AC_G$ , employing four, six and eight loudspeakers in the anechoic chamber with three reflective panels.

Consequently, the  $AC_G$  also present the same conclusions. The  $L = 8$  presents the most anechoic-like behaviour.

### Summary of Reflective surfaces

The aim of Complexity N°3 and Complexity N°4 was to evaluate the capabilities of the headrest-system to maintain the performance achieved in ideal free-field conditions and in the presence of reflective panels.

Contrary to the free-field conditions, when there is a strong reflection close to the  $DZ$ , the addition of the outer sources is ensuring that the  $AC$  is above 30 dB. Additionally, the system is capable of breaking its natural response imposed by the target pressure and provide an  $AC_G$  of more than 10 dB whilst maintaining an  $E_{rr}$  of the order of  $10 \times 10^{-4}$ , an  $E_{ff} \leq 3$  dB and a realistic Cond.

Regarding the resulting filters, it is possible to observe that in all the cases the sources in charge of the reproduction are almost ideally the FRF of the  $\delta[n]$  functions. In spite of obtaining more complicated cancellation filters the more complex the environment gets, especially when the wavelength starts being comparable to the elements in the system, the complex responses are only obtained in the cancellation filters, *i.e.* the reproduction should not be altered (or mildly altered) by their more complex shapes (this aspect is re-addressed in 3.5).

As a “bonus track” the presence of a reflective panel acting as ceiling was considered. It was observed that the  $L = 8$  system was capable of maintaining the performance in spite of adding a reflection “outside the plane” of the system.

### 3.4 Car cabin

The car cabin is assumed to be the most-complex acoustic environment of this Ph.D. Therefore, the prototype used in the laboratory conditions (Sec. 3.2.2) was introduced in a Peugeot 3008 SUV-type car (see Fig. 3.32).

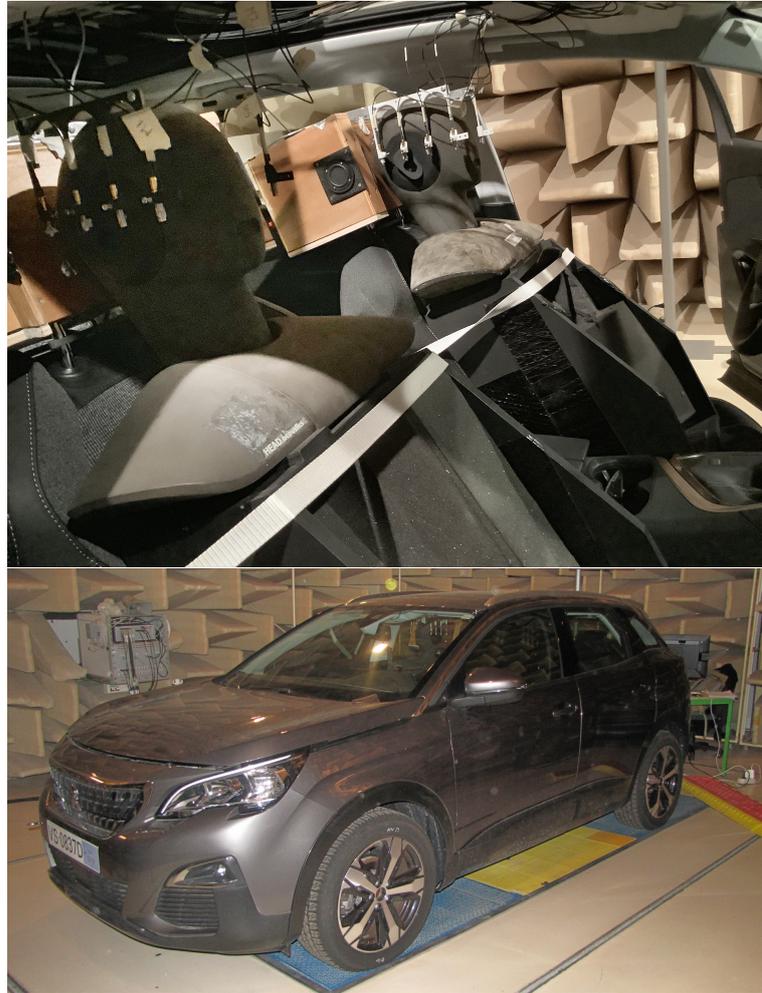


Figure 3.32: Experimental setup. Prototype in a car cabin with microphone array (above), car in semi-anechoic chamber (below).

For this experiment, the dimensions of the system are the same as used in laboratory conditions (Fig. 3.9). Both seats are separated by 73 cm, the same dummies, seats and headrests are used.

Lastly, it is important to mention that both seats located in the centre of the rails and the windows, trunk and sunroof are closed.

### 3.4.1 Complexity №5: Prototype in a cabin

Therefore, one proceeded to evaluate the resulting performance in the car cabin (Fig. 3.33).

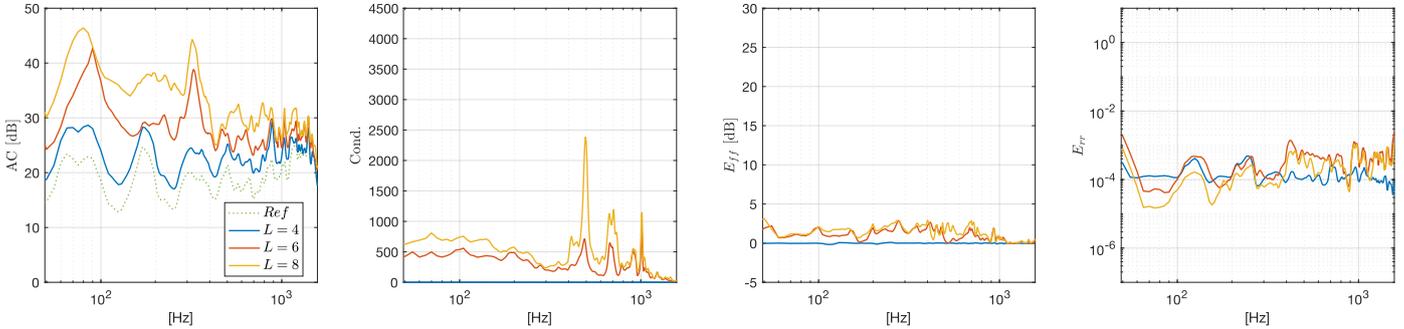


Figure 3.33: Resulting metrics employing four, six and eight loudspeakers in a car cabin.

All the tools/concepts used and developed to analyse the laboratory conditions can be directly extrapolated to study the complex car-cabin environment.

Unlike the more ideal scenarios, the car cabin's  $RAC$  presents a more oscillating behaviour due to the more complex acoustic conditions of the car-cabin. As reviewed, the natural response is still ruled by how the target pressure is "seen" in the  $DZ$  but now, it is no longer mirror sources interacting with the interference of  $P_R$  and  $P_L$  but their interaction with the modal response of the cabin; *i.e.* how these sources are driving the modes of the cabin and where the  $DZ$  is located, in a nodal line, in a maximum of pressure, etc.

As in laboratory conditions, the  $L = 4$  configuration is still mimicking the natural response of the system described by the  $Ref$ . The  $L = 6$  system, is capable of breaking the natural response and achieving a larger  $AC$ , but, as the system is more complicated it is not capable of maintaining a constant performance. Lastly, the  $L = 8$  system is presenting the larger performance, achieving an  $AC \approx 30$  dB for the complete bandwidth except for some drops at particular frequencies.

Regarding the reproduction error, in spite of maintaining similar levels the  $L = 6$  and  $L = 8$  for the first analysis the  $L = 4$  presents a larger reproduction error. This might be due to the fact that as an enclosure is being used, any energy introduced in the environment by any source arrives to the  $BZ$ .

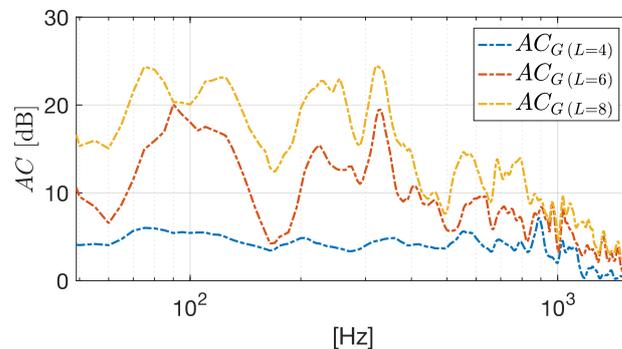


Figure 3.34: Resulting  $AC_G$ , employing four, six and eight loudspeakers in a car cabin.

Lastly, when analysing the  $AC_G$  in Fig. 3.34 a familiar behaviour can be seen. As it happened in all the cases studied before, the  $L = 4$  system is achieving a relatively constant of  $\simeq 4$  dB. Moreover, the  $L = 6$  configuration is presenting a larger  $AC_G$  but, as seen with the reflective panels (Fig. 3.27) it is still strongly

affected by the environment. When the  $L = 8$  system is being considered, the  $AC_G$  is outperforming the other configurations arriving to hold an improvement of  $\approx 10$  dB or more up to 900 Hz.

It is important to mention that, this general drop in performance happens around the typical Schroeder frequency of a car compartment (around 300 – 440 Hz [61, 62]).

In spite of the fact that the system was designed in free-field conditions (no reverberation, no modal response) its source distribution shows to be capable of controlling the field inside the car cabin. Below 400 Hz, the source distribution is capable of coupling with the acoustic modes of the enclosure and achieve levels of  $AC_G$  of the order of 15 dB. Above this frequency, where the modal density is high, the system is still capable of improving the performance, providing an  $AC_G$  of the order of 10 dB.

One proceeded to evaluate the resulting filters for the  $L = 8$  configuration in a car cabin.

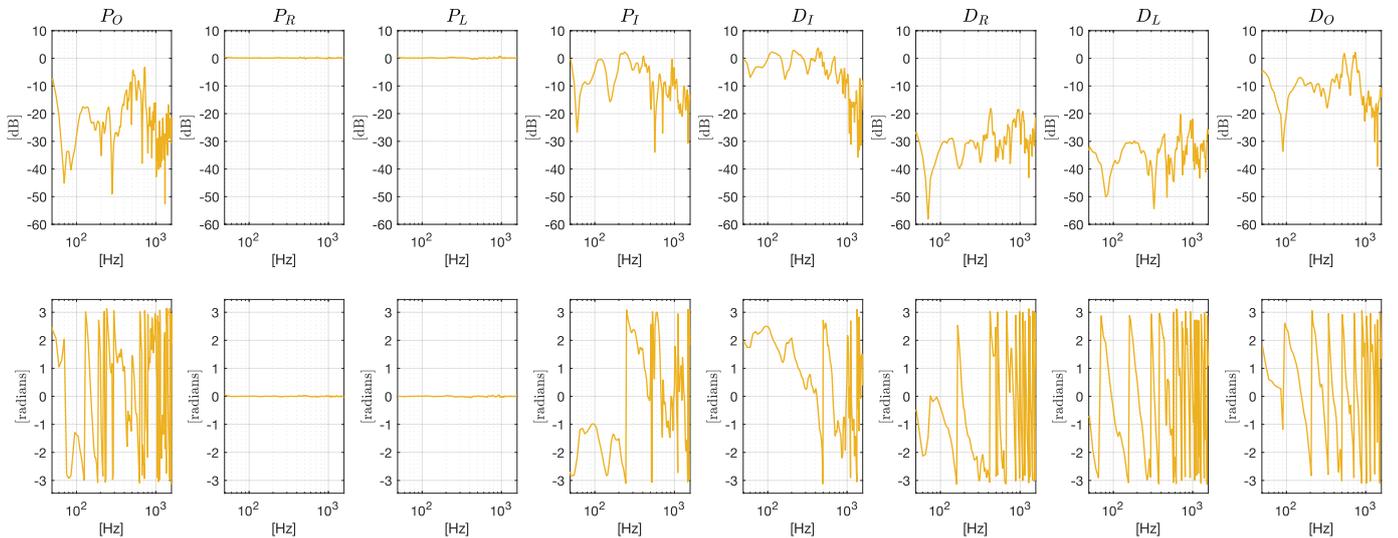


Figure 3.35: Resulting filters, eight loudspeakers in a car cabin.

When analysing the resulting filters in a car cabin, the first thing one notices is the reproduction sources and their almost perfect  $\delta[n]$  FRF. In addition,  $D_R$  and  $D_L$  are as in all the cases strongly attenuated. On the other hand, due to the complexity of the acoustic environment,  $P_I$ ,  $D_I$  and  $D_O$  are all strongly contributing to the cancellation part of the problem, being all of them driven with a very small attenuation. Lastly,  $P_O$  seems to be mostly contributing to the system in a narrow band between 400 Hz and 750 Hz.

The fact that the system is capable of improving the performance even above  $\simeq 400$  Hz is not only the effect of the near-field sources, but more of a joint-contribution of all the sources. Moreover, one can observe that the cancellation filters (especially  $D_I$  and  $D_O$ ) are also contributing significantly to the systems performance above this frequency (see Fig. 3.35). More evidence of this statement can also be seen in the differences between the  $AC_G$  above  $\simeq 400$  Hz when incorporating the extra sources (see Fig. 3.34).

On the other hand, it is important to highlight the fact that none of the resulting filters presented an amplification larger than 4 dB (see Fig. 3.35). This allows one to assume that the system is working in the linear regime of the loudspeakers.

### 3.4.1.1 The resulting pressure and the acoustic contrast

As in free-field conditions (Sec. 3.2.2) one proceeded to analyse the resulting pressure fields in the controlled zones for four different frequencies (Figs. 3.36 and 3.37).

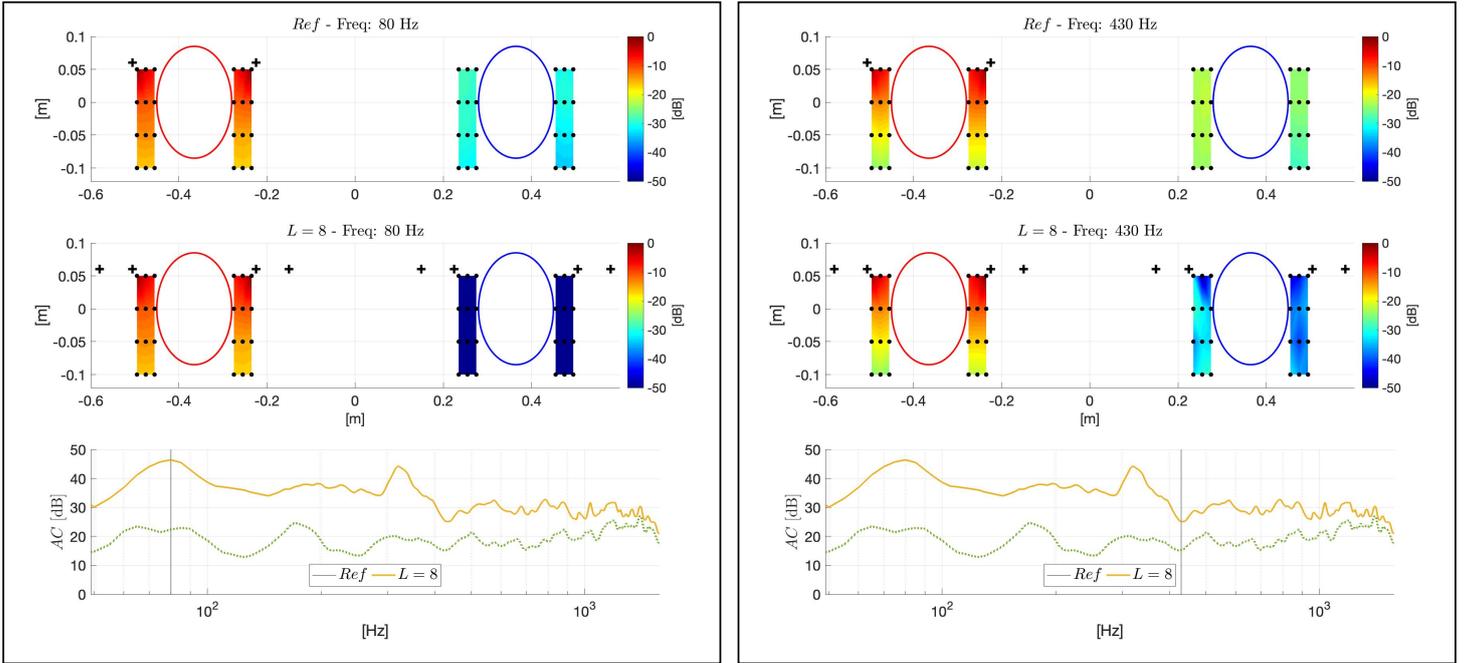


Figure 3.36: Normalised pressure for the  $L = 8$  and  $Ref$  configurations in a car cabin: At 80 Hz (left) and at 430 Hz (right).

Firstly, one could argue that at 80 Hz the loudspeakers are barely generating any SPL (resonance frequency around 145 Hz see App. C), this is a valid remark, but as the sources are placed in close proximity to the zones they are still capable of generating sufficient energy to control the fields (otherwise there would not be any improvement and the measurements would be noisier). In addition, looking at Fig. 3.36 (left) one can see evidence that at this frequency the natural response of the system is attenuating in the  $DZ \geq 30$  dB for the  $Ref$  (probably a longitudinal mode of the cabin), hence the system at this frequency achieves levels of attenuation of  $\simeq 50$  dB.

On the other hand, at 430 Hz there is an important loss of  $AC$  (see Fig. 3.36 right) which is the consequence of two effects happening simultaneously that can be seen in the resulting pressure field. Firstly, the natural response of the system is not so beneficial (its attenuation is  $\approx 25$  dB while for 80 Hz one had more than 30 dB). Secondly, it is possible to observe in the  $BZ$  that there is a steeper than usual reduction of the pressure in the outer microphones (at  $-0.1$  m  $\geq 25$  dB).

Additionally, one proceeded to evaluate the system's achieved  $AC$  for higher frequencies close to the bandwidth's limit (Fig. 3.37).

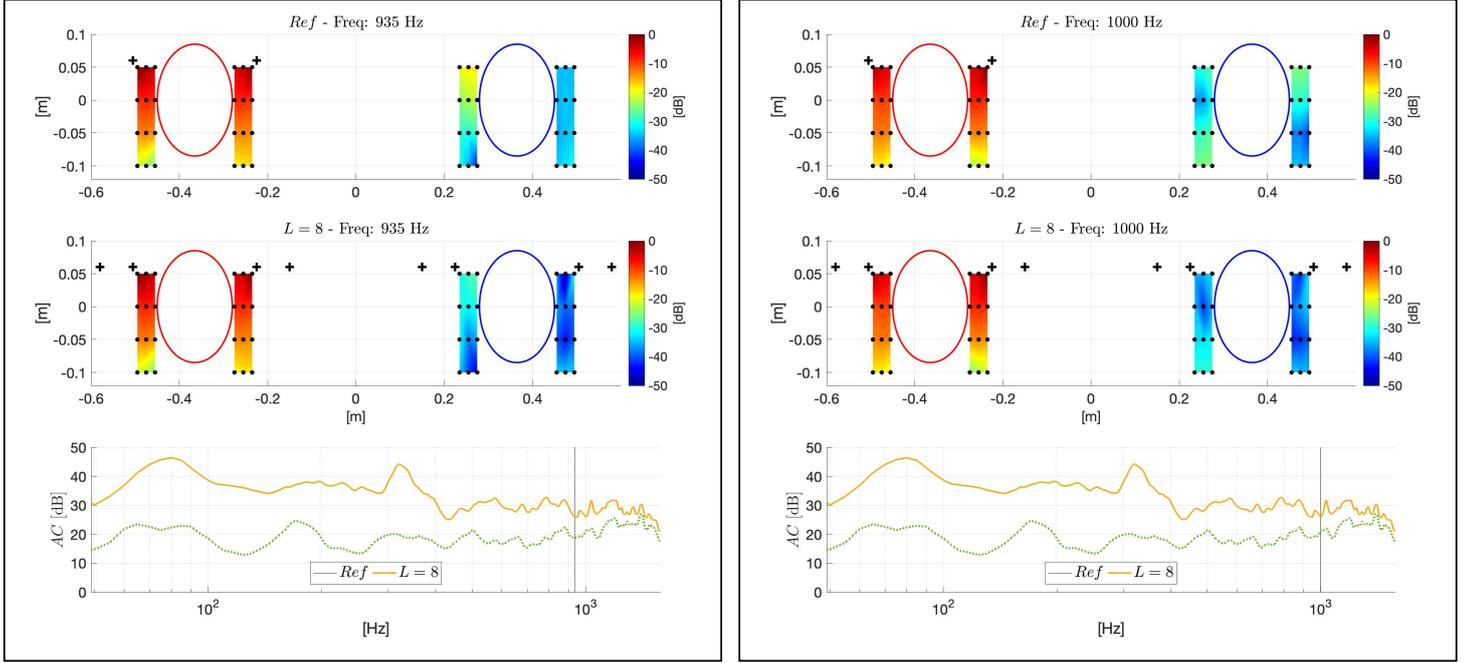


Figure 3.37: Normalised pressure for the  $L = 8$  and  $Ref$  configurations in a car cabin: At 935 Hz (left) and at 1000 Hz (right).

As the wavelength becomes more comparable to the size of the elements in the car cabin and the modal density becomes higher it becomes harder for the system to achieve high levels of attenuations resulting in a more uncontrolled pressure over the  $DZ$  (above  $\approx 900$  Hz). Still it is possible to observe some improvement (compared to the  $Ref$ 's attenuation) but the system becomes less reliable.

In this thesis, the  $AC$  and  $E_{rr}$  are used as the metrics of choice because they provide global information regarding the zones. However, as it can be seen in the resulting pressure fields (Figs. 3.36 and 3.37), as the modal density becomes higher, the shape of the resulting pressure in the zones becomes more complex. Even in the  $BZ$ , for the microphones the furthest from the loudspeakers, there are some contributions of the environment

Therefore, one could wonder if one could improve the performance by reducing the size of the controlled zones. In other words, if the performance of the system is susceptible to the size of the zones. If so, how does this improvement in terms of  $AC$ ,  $E_{rr}$  translates in terms of the resulting pressure fields.

To answer this questions, one proceeded to compare the system studied up to this point (15 cm long zones sampled 48 microphones) with a smaller zones system removing the microphone line the furthest from the loudspeakers (10 cm long zones sampled with 36 microphones) to maintain a realistic controlled area.

In order to test this hypothesis, the four different configurations listed in Table 3.3 are analysed.

	Zones for calculation	Zones for evaluation
a)	15 cm zones ( $K = 48$ )	15 cm zones ( $K = 48$ )
b)	15 cm zones ( $K = 48$ )	10 cm zones ( $K = 36$ )
c)	10 cm zones ( $K = 36$ )	15 cm zones ( $K = 48$ )
d)	10 cm zones ( $K = 36$ )	10 cm zones ( $K = 36$ )

Table 3.3: Cases to evaluate the effect of reducing the size of the zones.

The configurations in Table 3.3 are chosen to observe both the sensitivity of the metric (one zone size

used for calculation evaluated over different zones-size) and also to analyse the results of the optimisation over a smaller zone if it improves the performance (different zone sizes evaluated over one zone-size).

Therefore, one proceeded to plot the  $AC$ ,  $E_{rr}$ ,  $E_{ff}$  and condition number for the four configurations.

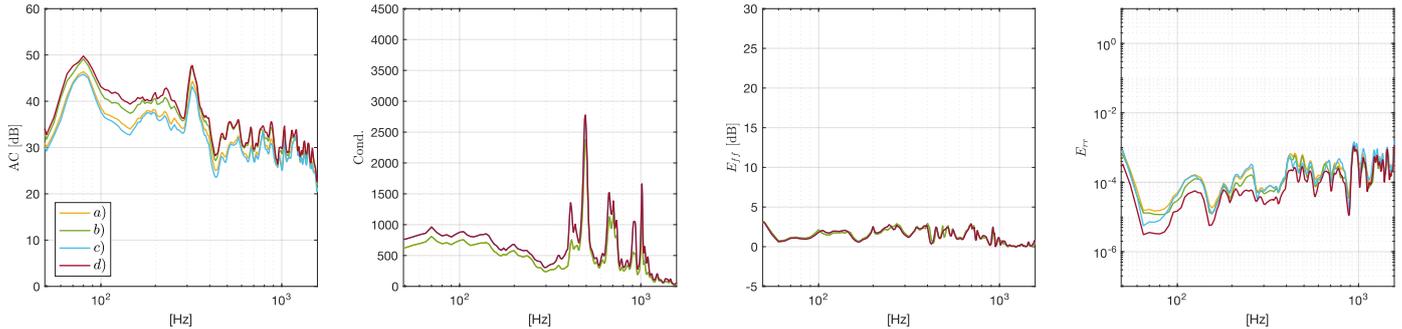


Figure 3.38: Resulting metrics, employing eight loudspeakers in a car cabin comparing configurations listed in Table 3.3.

Firstly, in spite of the fact that configurations  $a$ ) and  $b$ ) employ the same set of filters (both calculated over 15 cm zones), the fact that it is being evaluated over different zones resulted in a difference of  $\simeq 4$  dB of  $AC$  (see Fig. 3.38). The same behaviour is observed when comparing configurations  $c$ ) and  $d$ ) but with an even larger difference. Hence, despite using the same set of filters, *i.e.* perceptually there is no difference, the  $AC$  is strongly affected by the size of the control zones. Therefore, although the  $AC$  is an extremely useful metric, it cannot be used to compare among different zone-sizes.

Moreover, when comparing the resulting  $AC$  of configurations  $b$ ) and  $d$ ) (calculated over different sizes, but both evaluated in the same zone-size) one can observe  $d$ ) presents a slightly better  $AC$  for the frequencies below 200 Hz.

Regarding the  $E_{rr}$  (in Fig. 3.38),  $a$ ),  $b$ ) and  $c$ ) resulted in very similar reproduction errors despite the modification of the zone sizes. However, configuration  $d$ ) (calculated and evaluated over the smaller zone) presents the lowest  $E_{rr}$  of all the studied configurations. On the other hand, it is possible to observe that the reduction of microphones resulted in a slight growth of Condition number, whilst the resulting  $E_{ff}$  is almost identical in all four configurations.

Lastly, one proceeded to plot the resulting pressures (as in Fig. 3.37) when the last row of microphones was removed both for calculating the filters as well as for evaluating the performance (configuration  $d$ ) in Table 3.3).

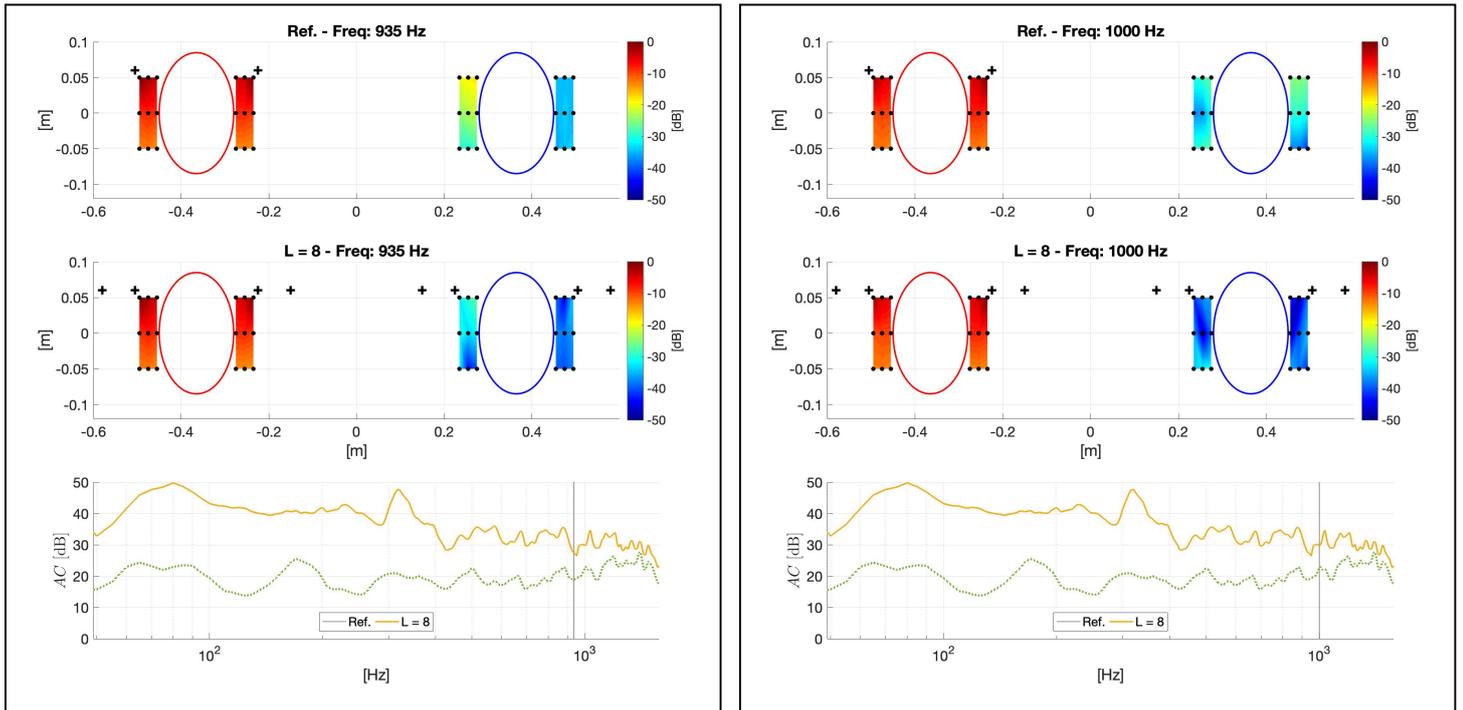


Figure 3.39: Normalised pressure for the  $L = 8$  and  $Ref$  configurations in a car cabin using 36 microphones: At 935 Hz (left) and at 1000 Hz (right).

When one compares Fig. 3.37 (configuration *a*) and Fig. 3.39 (configuration *d*), it is not possible to observe any noticeable differences between the two configurations, besides the slightly  $AC$  enhancement and the missing microphones.

Lastly, in this section one studied the resulting pressures in the car cabin. The system was still capable of improving the performance up to 1 KHz. The higher one goes in frequency (with a noticeable break at  $\simeq 430$  Hz), the modal density of the car cabin becomes higher and the resulting pressure fields become more difficult for the system to control.

On the other hand, one added evidence to the statement (in Sec. 2.5) that one needs to be very cautious when incorporating conclusions from other studies. It was observed that even applying the same filters, depending on the size of the zones one obtains an  $AC$  boost up to 7 dB. Therefore, it is fundamental to use identical conditions to compare between different systems, methods, etc.

Due to the the fact that the outer line of microphones is probably never going to be reached by the listener and its information might not be representative to the perception of the PSZ problem, from this point forward one is going to reduce the size of the zones to  $4\text{ cm} \times 10\text{ cm}$  on each ear, a total of 36 microphones.

### 3.4.1.2 “Mono” vs. “Stereo”

One of the original requirements of the system (Chapter 1) is to be compatible with other sound spatialisation technologies. Up to this point, due to the  $\delta[n]$  reproduction filters, as well as the low  $E_{rr}$ , it was assumed that the simultaneous use of a different technology would not present major complications. However, these experiments were performed assuming the reproduction of a “mono signal”, which is too restrictive to implement more complex spatialisation technologies.

This section is introduced as a brief extension of the previous experiments when different programs are reproduced on the right and left channels of the  $BZ$ , *e.g.* when the passenger is the  $BZ$ , one program reproduced from  $P_R$  and a different one from  $P_L$ .

Moreover, it is necessary to mention that, for each  $BZ/DZ$  definition, one needs to calculate two sets of optimal filters and “exploit” the linearity assumption of the CO methods (see Sec. 2.2.3).

Therefore, in order to analyse the “stereo” performance of the system, four different cases are compared:

- *Mono*: The desired pressure is the resulting pressure when both  $P_R$  and  $P_L$  are driven in phase.
- $P_R$ : The desired pressure is the resulting pressure when only  $P_R$  is driven.
- $P_L$ : The desired pressure is the resulting pressure when only  $P_L$  is driven.
- $P_R + P_L$ : The linear superposition of the resulting filters when only  $P_R$  is driven and when only  $P_L$  is driven.

One proceeded to calculate the resulting metrics for the four aforementioned cases.

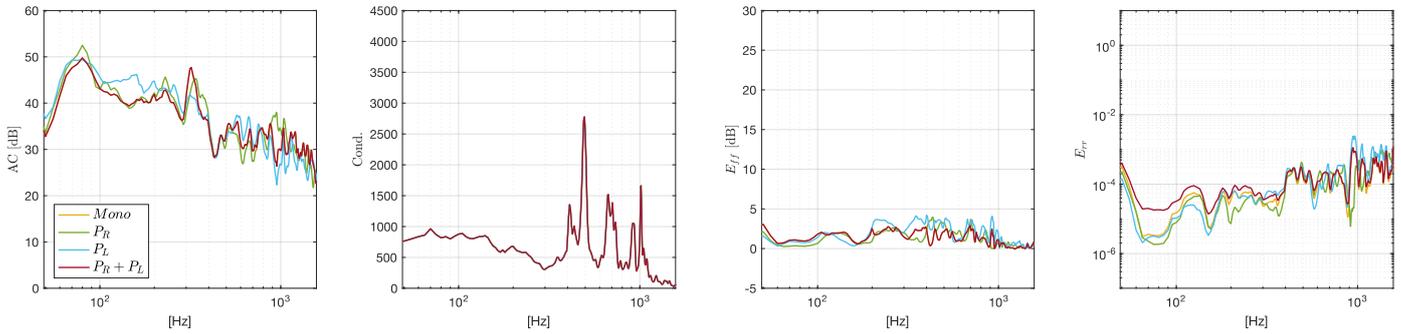


Figure 3.40: Resulting metrics, employing eight loudspeakers in a car cabin when generating a Mono signal (yellow), a signal only from the right (green), a signal only from the left (cyan) and a stereo signal (dark red).

Firstly, it is important to remember that the definition of the desired does not have an influence on the condition number of the system (the inverted matrix does not change see Sec. 2.3.3.3), hence, all four cases result in identical condition number. In addition, the  $E_{ff}$  is slightly different between all four cases, but still in all the cases is below 5 dB.

Furthermore, it is possible to observe that the resulting  $AC$ , when using the *Mono* filters and the  $P_R + P_L$  filters, is identical. Additionally, the other two cases present comparable performances except for a deterioration above 950 Hz when only  $P_L$  is used. Regarding the  $E_{err}$ , all four configurations presented the same order of performance. There is a small deterioration with the  $P_R + P_L$  configuration below 100 Hz, but still the order of magnitude is very small.

In spite of the encouraging results, it is important to highlight the limitations of this study. Firstly, in theory any spatialisation technology can be applied, as long as the linearity assumption is maintained. If the spatialisation chosen produces the sources to work in their non-linear regime it will affect the PSZ performance [142].

Furthermore, one could extend the PSZ formulation to incorporate the spatialisation technology within the optimisation. For example, one technology that presents particular interest is the implementation of binaural technologies using crosstalk cancellation [143, 144]. Moreover, by defining only one ear as the  $BZ$  and the other one as the  $DZ$ , one “incorporates” the inter-aural crosstalk in the optimisation (as studied for example in Refs. [10, 20] among others).

In addition, another example of the incorporation of the spatialisation of sound in the PSZ was published by Coleman *et al.* in 2014 [145]. In this paper, the authors studied the use of a 60 loudspeaker system around the zones to synthesize two plane waves to approximate a stereophonic reproduction. Moreover, the fact that in the system studied in this thesis, one loudspeaker is placed on each side of the head, and that different programs are being reproduced, it cannot be considered as a stereophonic system. A stereophonic format requires two loudspeakers at  $\pm 30^\circ$  in front of the listener.

An extensive study on accurate reproduction of perceived level and timbre to ensure the compatibility of a similar system with binaural recording can be found in Ref. [146].

Although in this thesis one is going to maintain the study of a mono signal (attempting to maintain the reproduction in the *BZ* as “transparent” as possible), it is important to consider in future works the incorporation of the spatialisation technologies. Their implementation should be evaluated as a system in series (exploiting the linearity assumption) as well as incorporating the spacialisation principle in the PSZ cost function (modifying the definition of the control zones and/or the desired pressure accordingly).

### 3.5 Notes on the resulting filters: frequency or time formulation

Up to this point, a frequency domain formulation has been used to calculate the best filter combination. This formulation was chosen because it allowed to make an analysis solely focusing on the physical limitations of the system neglecting the effects of frequency band truncation, windowing, etc. However, this is a bit far from reality, as stated by Møller and Olsen in 2017 (...) *the envelope of the control filters is related to the perceived audio quality of the target sound field in the bright zone* [118], therefore if a transparent system is wanted, the time-domain response of the filters also becomes a parameter to take into consideration.

The study of the filter’s time-domain response -and their causality- has already been addressed in the literature. On the one hand, Møller *et al.* [118] proposed to solve this issue incorporating the causality/time envelope as a constraint in the cost function. In spite of the cleverness of this approach, the fact that the filter’s time-shape or causality is added as a constraint in the optimisation can result in a modification of the overall-performance of the system because the cost function is being changed. In this section, the “combined solution” (CS) is maintained as the cost function of choice with an equal weighting within its terms (error minimisation in the *BZ* and an energy minimisation in the *DZ*) without the addition of any extra constraints (see Sec. 2.3.3.4).

The combined solution was compared when formulated in the time domain and in the frequency domain in an anechoic environment in Ref. [125] or, as labelled in this report, Complexity №2 (Sec. 3.2.2). Hence, this sub-chapter is presented as an extension of the aforementioned paper using the results obtained in a car cabin or Complexity №5 (Sec. 3.4).

As in Ref. [125], two formulations are studied to calculate the optimal filters using the CS, one in the frequency domain (PM) and one in the time domain (BACCPM). The main difference between these formulations is that, while the frequency domain calculates each frequency as individual optimisation problems, the time domain formulation optimises the whole available bandwidth in a single step [97]. Furthermore, in Ref. [125] it was observed that the time domain formulation presents a more computational-demanding algorithm which resulted in a 500 times slower calculation of the filters than its frequency domain counterpart.

#### 3.5.1 Time domain

Firstly, one is going to analyse the definitions of the time-domain formulation in order to make the comparison between both methods.

As mentioned in Chapter 1, the upper frequency limit on this report is  $\approx 1$  kHz, therefore it is necessary to truncate the results accordingly in order to obtain the filter’s IRs, therefore the calculations are performed using a sampling frequency  $F_s = 2450$  Hz, hence the Nyquist frequency is 1225 Hz. Moreover, to obtain the IRs of the resulting filters from the frequency domain optimisation, an inverse Fourier transform is applied. Therefore, due to the frequency truncation and the 5 Hz resolution, the resulting FIR are 490 coefficient filters, hence to make an equal comparison, the time domain formulation is calculated with the same length.

In order to avoid errors in “very low frequency” (below 50 Hz) due to the loudspeaker’s response, a high-pass filter is applied to the resulting filters: a 301 coefficients FIR filters with a cut-off frequency of

36.75 Hz (see Fig. 3.41).

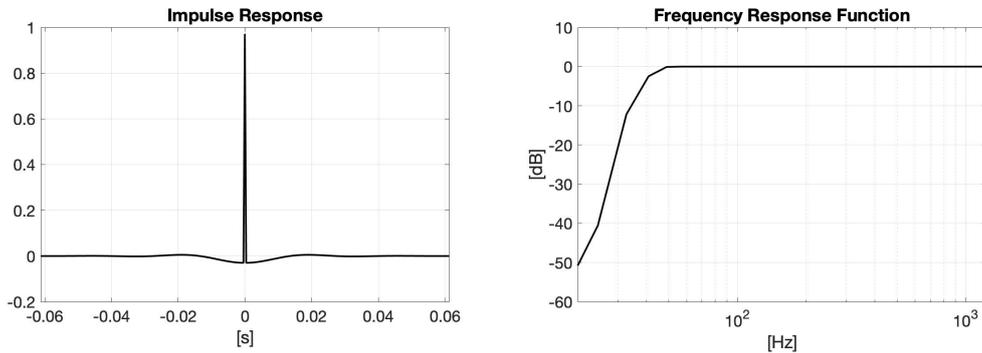


Figure 3.41: Filtered used to remove noise below 50 Hz: IR (left), FRF (right).

It is possible to observe in Fig. 3.41 that the filter is only targeted to remove any possible noise in the calculation and to keep a flat response above 50 Hz.

### The modelling delay

In Ref. [97], Galvez *et.al.* mentioned among the advantages of the time domain formulation the possibility to choose the amount of modelling delay desired to process the array. Therefore, one proceeded to study -similarly to the study in Ref. [97]- the effect of the modelling delay in the car-application. For this study, the length of the filters was fixed to  $M = 490$  samples.

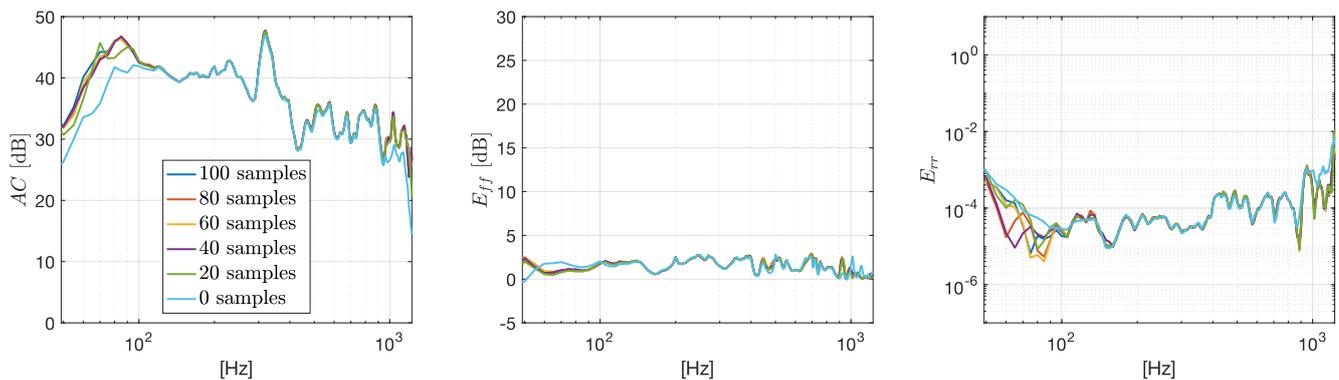


Figure 3.42: Resulting metrics, employing eight loudspeakers in a car cabin for different delays.

When analysing Fig. 3.42, it is possible to observe that the resulting  $AC$  and the  $E_{rr}$  between  $\approx 100$  Hz and  $\approx 900$  Hz are not affected by the modelling delay.

However, above 900 Hz it is necessary to have a small modelling delay  $\geq 20$  samples (around 8.2 ms) to achieve a more stable performance. In addition, below 100 Hz, a similar behaviour is found but slightly more chaotic, the  $E_{rr}$  keeps changing for the different delays and the  $AC$  seems to achieve a relatively constant value for delays greater than 20 samples. Nevertheless, the values achieved at these frequencies is acceptable for every modelling delay above 20 samples.

It is also necessary to evaluate the effect of the modelling delay in the resulting filters, for display purposes, all the modelling delays are compensated after the filter calculation.

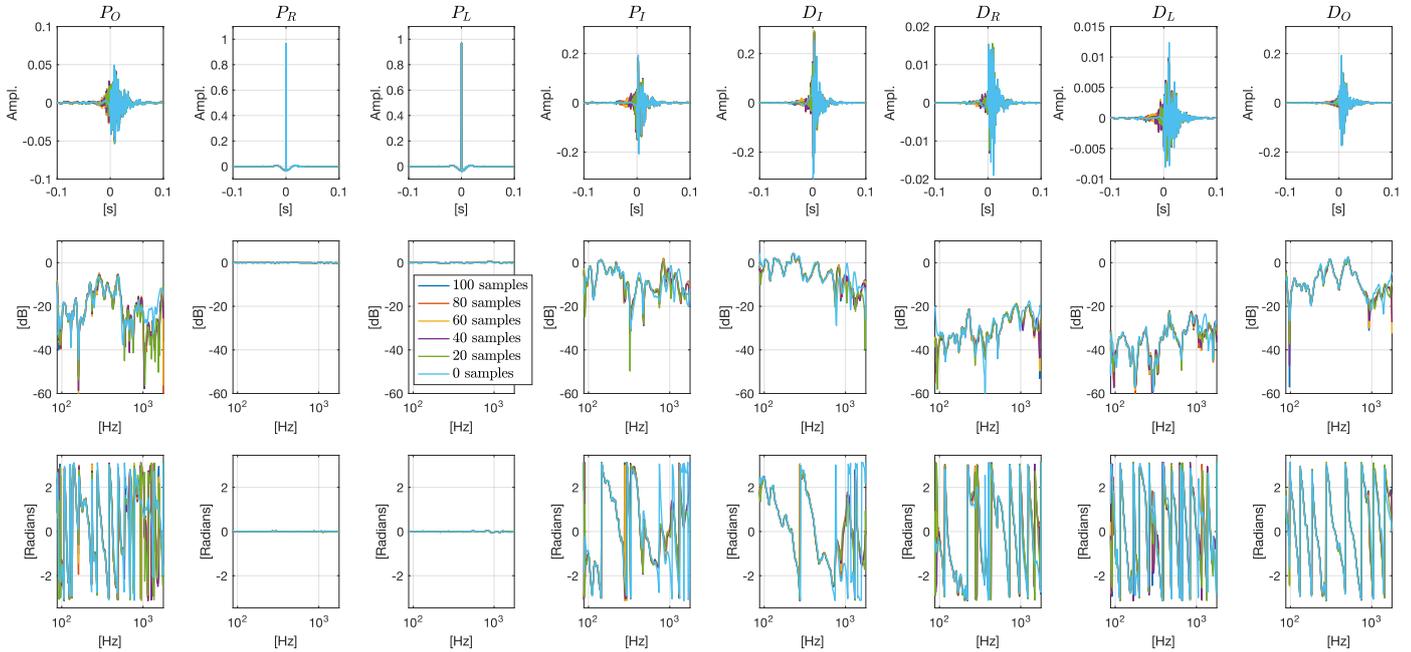


Figure 3.43: Resulting Filters, for the different modelling delays. To simplify the comparison between the different resulting filters, all the modelling delays are compensated after the filter calculation.

In order to be able to compare between the different modelling delays, all the resulting IRs are centred and the modelling delays compensated to synchronise all the filters.

Firstly, it is possible to observe that the IRs of the reproduction filters are an almost identical copy of the high-pass filter used to remove the very-low frequency noise (see Fig. 3.41), *i.e.* the filters associated to the sources  $P_R$  and  $P_L$  result in almost perfect  $\delta[n]$  functions (convoluted with the high-pass filter).

Lastly, the more graphical effect of the modelling delay can be seen in the resulting IRs of the cancellation filters, especially in  $P_I$ . It is possible to observe that when no modelling delay is applied, there is almost no pre-ringing and as the modelling delay grows, how the pre-ringing gets longer and longer.

There was a particular interest in Ref. [97] to maintain the modelling delay as small as possible in order to be able to synchronise the PSZ with a television, therefore the modelling delay was considered an important design factor in order to achieve image-sound synchronisation. However, in the car application this is not such a restriction due to the fact that there is no need to synchronise it with an external device. Nevertheless the modelling delay has a strong effect on the pre-ringing of the resulting IR, therefore, as a trade-off between performance and pre-ringing, in the following studies a delay of 20 samples ( $\approx 8.2$  ms or approximately a 25th of the IRs length).

### Length of the filters

As mentioned above, the length of the resulting FIR filters was fixed to fit the same “resolution” as the frequency domain formulation. However, the time domain formulation allows to fix the length of the resulting IRs without any extra tampering needed. Therefore, one calculated the resulting metrics for ten different filter lengths applying a fixed modelling delay of 20 samples (Fig. 3.44).

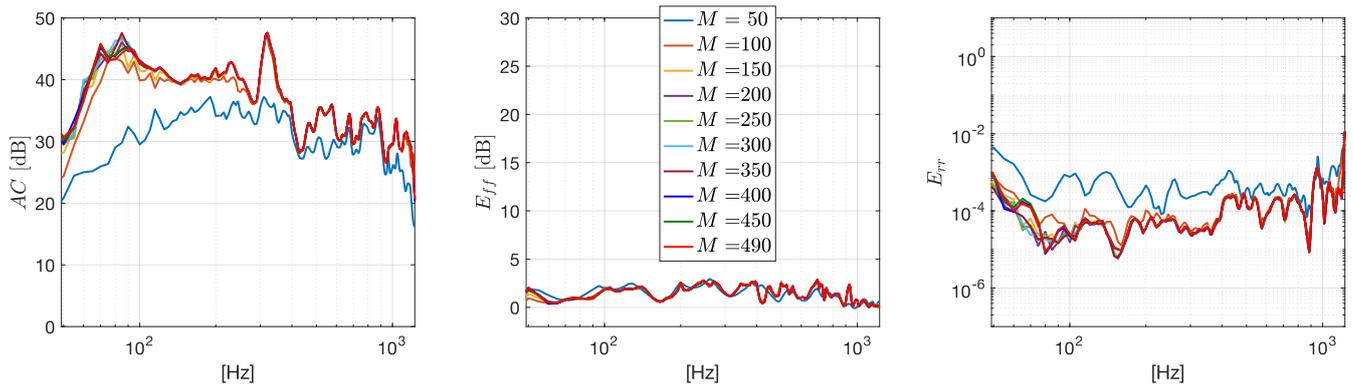


Figure 3.44: Results for different filter lengths with modelling delay fixed to 20 samples.

One can observe in Figs. 3.44 that after 200 coefficients, the number of coefficients has very little effect on the performance of the system, their effect mostly changes the performance below  $\approx 120$  Hz. Therefore, one could reduce the number of filter coefficients if necessary without having strong changes in the performance.

### 3.5.2 Time or Frequency?

As it was explained in Ref. [125], in order to compare the performance of both formulations it is necessary to make the algorithms optimise over the same conditions (at least as similar as possible), *i.e.* the resulting “frequency-resolution”, given by the length of the IRs in the time domain, needs to match the frequency step used in the frequency domain formulation, as well as the length of the calculated filters. This equivalence was recently re-addressed by Lee *et al.* in Ref. [147].

Therefore, both methods are used to calculate 490 coefficient filters, and a modelling delay of 20 samples is used for the calculation. One proceeded to evaluate the resulting metrics using both the PM and the BACCPM.

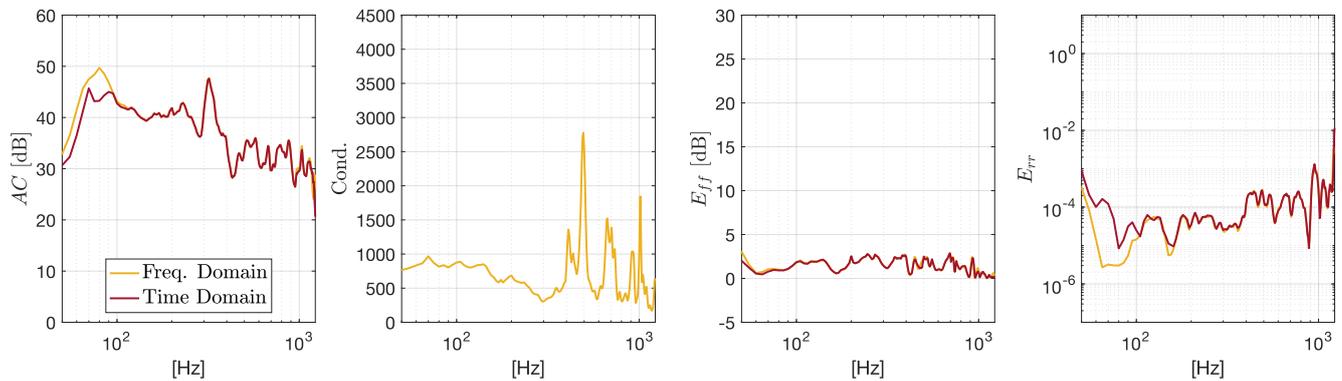


Figure 3.45: Results in a car cabin, time vs. frequency formulations.

As observed in Ref. [125], from 100 Hz to 1 KHz both formulations result in almost identical performances (Fig. 3.45). Below 100 Hz, the frequency domain outperforms the time domain formulation achieving slightly better  $AC$  and  $E_{rr}$ , but this difference is negligible.

It is important to take some time and discuss the condition number. As it was mentioned, as the time domain optimises the complete frequency band once, there will be only one condition number unlike the frequency domain that presents a frequency dependent conditioning. The condition number for the time domain was  $\text{Cond.} = 1.24 \times 10^6$ . Although presenting a more elevated condition number, none of the previous cases presented serious numerical complications, furthermore, in App. D one can find the resulting  $AC$  measured in a car cabin in an online manner. Therefore it is possible (and for the seek of maintaining a fair comparison) to move forward without regularisation.

Nevertheless, by analysing the Cond. curve in Fig. 3.45 it is possible to observe that regularisation might only be needed at some very specific frequencies (*e.g.* at 495 Hz and at 1015 Hz) where the condition number grew. This type of punctual regularisation can be easily performed in the frequency domain by monitoring frequency by frequency both the effort and the Cond. number and only adding regularisation if needed (however this will increase the computational requirements of the algorithm). The implementation of a frequency dependent regularisation in the time domain requires to known in advance the regularisation shape to define the FIR filter to regularise (see Sec. 2.3.3.4).

On the other hand, to be congruent with [125], the time taken by each formulation to obtain the filter's IRs -without regularisation- from the matrix of the system's IRs is given. The frequency domain took 7.2 ms to obtain the filter combination whilst the time domain took 12.6 s, more than 1000 times longer. The calculation time is expressed as the mean value after  $2 \times 10^4$  iterations. The code was implemented under MATLAB, under Windows 7 Enterprise, using an Intel Core(TM) i7-6820HQ CPU with a RAM memory of 16 GB and a 64-bit architecture.

Although this is not the case in this thesis, if one could bear to have some performance degradation one could reduce the computational time by using the conjugate gradient method proposed by Shi in Ref. [148].

### The resulting filters

Lastly, one proceeded to study the resulting filters for both formulations (Fig. 3.46).

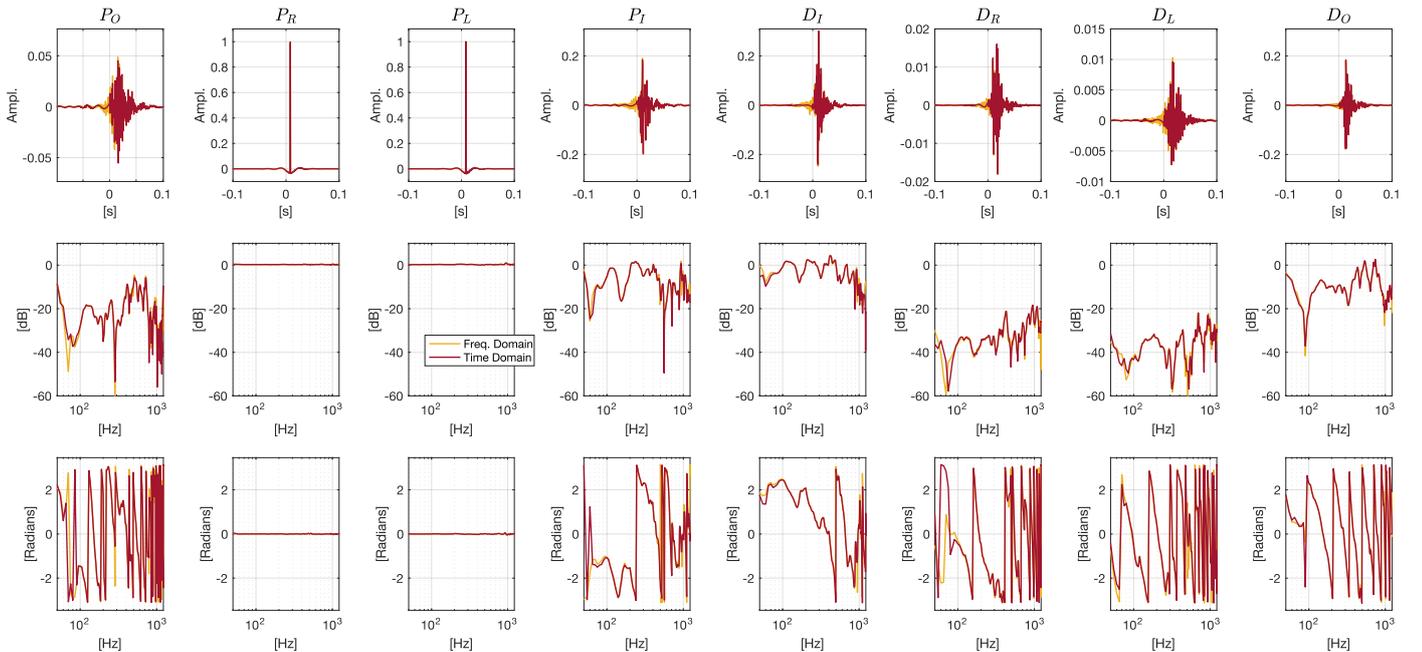


Figure 3.46: Resulting Filters, frequency (yellow) and time (red) domain formulations. Modelling delay removed from FRFs to simplify the analysis.

When analysing the frequency responses (both amplitude and phase in Fig. 3.46), it is possible to conclude that both methods achieve very similar performances. Nevertheless, it is possible to distinguish some differences for some particular frequencies.

To evaluate the impulse responses it is useful -once again- to separate between reproduction and cancellation filters. On the one hand, the reproduction filters are identical regardless the formulation, but they are no longer presenting the IR of a perfect  $\delta[n]$ . This due to the fact that a high-pass filter was used to avoid noise below 50 Hz and that a modelling delay of 8.2 ms is used; *i.e.* the reproduction filters are now the convolution between a  $\delta[n]$ , and the delayed high-pass filter IR.

On the other hand, the main difference can be seen in the cancellation filters, where -as observed by [97] the time domain formulation present smaller pre-ringing. Therefore if the frequency domain filters is chosen an extra delay needs to be applied to use the filters in real-time ( $\approx 50$  ms should suffice).

### Summary of filters

This section focused on the comparison between the two main formulations used to solve the combined-solution cost function. In addition, it was presented as an extension of the results obtained in [125] but in a car cabin environment (Complexity №5).

As seen in anechoic conditions (see Ref. [125]), although the time domain presents a slight deterioration of the performance below 100 Hz -both in  $AC$  and  $E_{rr}$ - both formulations are just as efficient in finding the best filter combination.

It is important to bring up one of the main conclusion from Ref. [125]. In spite of the fact that the time domain formulation presents IRs with smaller pre-ringing, the amount of computational power needed for this formulation is considerably larger, consequently it takes more time to calculate the resulting filters.

Therefore, if one had to answer the question, which formulation is better? or more properly: which formulation provides the best trade-off for the studied application? At this point there is no correct answer. As reviewed in Ref. [125], it highly depends on the calculation limitations of the system. If the calculation-time/computational-cost is not a constraint, the time domain formulation achieves almost the same performance with “nicer-IRs” as mentioned in Ref. [97]. On the other hand, if the filters need to be calculated in the moment (or many times), the computational cost required by the time-domain formulation may be an impediment for the system.

In the next chapters, depending on the problem in hand, one is going to either choose a time domain or a frequency domain formulation and the main arguments for this always goes around the computational cost constraint.

## 3.6 Summary

The aim of this chapter was to design a PSZ system capable of achieving the desired performance in a car cabin. For this, five different levels of acoustic complexity were considered in order to design and test the system. All the experiments carried out in this chapter are summarised in Table 3.4. In the following table, one can find the mean  $AC$  and  $E_{rr}$  between 50 Hz and 1225 Hz in the five different acoustic conditions for different loudspeaker combinations.

Mean values [50 Hz - 1225 Hz] evaluated using $K = 36$								
Complexity	Conditions	$Ref$ [dB]	$L = 4$		$L = 6$		$L = 8$	
			$AC$ [dB]	$E_{rr} \times 10^{-4}$	$AC$ [dB]	$E_{rr} \times 10^{-4}$	$AC$ [dB]	$E_{rr} \times 10^{-4}$
№1	Monopoles Free-field	24.2	27.6	0.25	37.2	0.72	40.3	0.37
№2	Prototype Anechoic chamber	21.5	27.5	0.33	39.4	1.53	41.8	1.34
№3	Monopoles Two parallel baffles	23.3	26.8	0.73	34.9	2.18	39.1	0.57
№4	Prototype Two parallel panels	21.9	27.0	0.43	35.3	2.68	40.9	1.94
№5	Prototype Car cabin	19.7	24.5	2.12	29.4	2.94	34.7	1.69

Table 3.4: Mean  $AC$  and  $E_{rr}$  for the five different levels of acoustic complexity for different number of sources.

In spite of the frequency-dependent characteristic of the system, the global metrics introduced in Table 3.4 provide an interesting overview of the results under static conditions.

Firstly, in Sec. 3.1, one defined the zones to control and established as starting point of the study the use of loudspeakers in headrest. This choice is due to its popularity in the car industry and the fact that it allows to extend the bandwidth of the system, improve the decoupling with the environment and provide an  $AC$  of  $\simeq 20$  dB without any processing (see Table 3.4).

In Sec. 3.2.1, the system was tested using theoretical monopole sources in free-field conditions (complexity N°1) where the addition of more sources was decided having in check both terms of the studied cost function (see Eqs. 2.30 and 2.31). Furthermore, in Sec. 3.2.2, a prototype was designed and constructed in order to validate these results experimentally. A series of measurements were carried out in an anechoic chamber employing two seats, custom made headrests and two dummy-head and torso as listeners (complexity N°2). It was found in the anechoic chamber (complexity N°2) that the strongest contribution to the system's performance was obtained by incorporating sources between the two zones ( $L = 6$ ) and that the contribution of sources in the exterior part of the problem did not contribute much to the system's performance (see Table 3.4).

As one step closer to the car application, in Sec. 3.3, one proceeded to repeat the experiment using controlled reflections. The system was evaluated using a simple first-order reflection model (complexity N°3) as well as measurements in an anechoic chamber incorporating reflective panels (complexity N°4). Although the system was still capable of maintaining the wanted performance, it was observed that in the presence of reflections, the contribution of the outer sources played a key role in maintaining the  $AC$  of the system as well as a low  $E_{rr}$  (see Table 3.4).

Lastly, in Sec. 3.4, the prototype was later introduced in a car cabin (complexity N°5) where the eight loudspeaker system was capable of maintaining 30 dB of  $AC$  whilst maintaining a low reproduction error for the studied bandwidth. In the car cabin, the  $L = 8$  system provided a mean  $AC$  of 34 dB and showed the lowest mean  $E_{rr}$  of all the studied configurations (see Table 3.4).

Throughout the different levels of complexity, not only the  $AC$  and  $E_{rr}$  were tested. It was also analysed that the  $E_{ff}$  demanded by the system is in all the experiments below 5 dB and the problem did not show any ill-conditioning problems (validated in Appendix D). Additionally, the resulting pressure as well as the resulting filters (both frequency and time responses) were thoroughly analysed. A dedicated section was introduced (Sec. 3.5) comparing both the frequency and time formulations, their different parameters, their resulting filters (in the time domain) and their resulting performances.

### The balance of the chapter

The main contributions of the chapter are listed below.

- 1) Design and validation of a feasible PSZ system, for a car application, capable of achieving  $\simeq 30$  dB of acoustic contrast whilst maintaining a low reproduction error employing only eight channels.
  - Patent pending: N°FR1907364.
- 2) Comparison between frequency and time formulations of PM formulation for PSZ under identical conditions.
  - Presented at the 147<sup>th</sup> AES Convention proceeding listed in Ref. [125].

# Chapter 4

## Variable acoustics

Up to this point, a system capable of fulfilling the original constraints was designed and tested in five different levels of acoustic complexity, from laboratory conditions (Secs. 3.2 and 3.3) to a real car cabin (Sec. 3.4). In all these scenarios the system proved to be able to improve the natural acoustic contrast on average by 10 dB for the frequency band of interest whilst maintaining a very low deterioration of the reproduced signal.

Moreover, one of the strongest assumptions made in the previous chapters was that the acoustical conditions were invariant, hence, it was possible to calculate the filters from the measured IRs -using a deterministic approach see Sec. 2.3.1- and consider that they will be the best solution possible for the system. However, a car cabin is an acoustic environment that changes over time, so it makes one wonder, **what would happen if there is an event that alters the acoustical conditions of the overall problem? is it necessary to re-measure all the IRs and re-calculate the filters? is the system still capable of generating PSZs in the new acoustic conditions?**

Therefore, the aim of this chapter is to answer all these questions from the “deterministic-approach” point of view. Firstly, one proceeded to extend the literature review (from chapter 2) by incorporating studies involving modifications in the acoustic conditions. Moreover, a review on which of these modifications can alter the performance of the designed system is introduced.

Secondly, an experimental procedure is explained and implemented to study the sensitivity of the PSZ system to these modifications. Lastly, one evaluates the possibility of using a deterministic approach, in a variable acoustic situation, to then study its advantages and disadvantages.

### 4.1 Literature review extension: PSZ under variable acoustics

There are several studies reviewing the sensitivity of the system to variabilities in the acoustic environment; *i.e.* modifications of the IRs. The aim of this section is to review which modifications have already been studied using deterministic approaches, identify which ones particularly concern our application and review the proposed solutions.

It is important to mention that some of these changes have not necessarily been evaluated experimentally, some of the inaccuracies are “artificially” added emulating a physical modification.

- **Source FRFs inaccuracies:** Several authors studied the influence of inaccuracies in the loudspeaker’s FRFs in PSZ systems. In 2012, Elliott *et al.* [129] analysed the effect of modifying the sensitivity of one driver and its effect on the resulting *AC* mostly employing the ACC method. The solution proposed by Elliott is to implement a frequency-dependent effort regularisation term to maximise the *AC* in the presence of variations of  $\pm 1$  dB in the response of one driver.

Additionally, Park in 2013 [149] studied this phenomena by incorporating analytically both magnitude and phase variations in the TFs. This idea of artificially incorporating inaccuracies was later implemented by Cai *et al.* with the BACC-RV (see Table 2.1) and expanded by Zhu *et al.* in 2017 [150] and 2019 [131].

As mentioned in Sec. 2.3.3.4 most of these methods also propose to overcome the variability of the sources FRFs by adding regularisation to the different cost functions. Each author presented a different alternative in how to choose this regularisation parameters more efficiently.

On the other hand, in 2017 Ma *et al.* [151] studied the influence of the non-linear distortion of loudspeakers on PSZ systems which resulted in a considerable loss of performance, the experiments were performed in an anechoic chamber. In this paper, the authors proposed to add a regularisation term to reduce the non-linear distortion.

In this dissertation, due to the small effort -not exceeding 5 dB- and the fact that none of the resulting filters resulted in strong amplifications (see Sec. 3.4.1), it is possible to assume that the loudspeakers are working in their linear regime, therefore one can neglect the distortion due to non-linearities on the loudspeakers. On the other hand, product variability as well as loudspeaker wear are two main concerns that can alter the overall response of the system and can not be neglected in the studied application.

- **Uncertainties in the TFs due to source/zones positioning and environment:** In 2013 park *et al.* also reviewed the sensitivity of the ACC to magnitude and phase variation due to changes in the source positioning [149]. In addition, in 2012 Elliott *et al.* evaluated the effect of the diffuse field and also the effect of the the driver positioning [129].

From these previous studies it is possible to derive two main conclusions that can not be avoided in our system:

- Firstly, it is necessary to know precisely the positions of the sources and zones.
- Secondly, it is also necessary to know the environment’s contributions, both direct and diffuse field.

In this thesis (in Sec. 3.4), the system was evaluated in a car environment, however, the seats and the headrest were fixed in one position (centre of the seat-rails and the headrest in the bottom position). Additionally, both the zones as well as the sources are linked to the seat/headrest. Therefore, if one modifies either of these, it would change not only the geometry of the problem (distancing the zones) but also its interaction with the modal response of the cabin.

In this dissertation it is assumed that both the seat positioning as well as the headrest positioning are strongest changes in the acoustical characteristics of the system.

- **Passenger changes in a car:** In 2019, Cho *et al.* [152] studied the modification of the acoustic transfer functions due to different passenger configurations with/without front and rear passengers and their posture. They observed a considerable variation in the TF but their effect on a PSZ system was not reviewed.

All the aforementioned studies focused their research in studying the effect of TF inaccuracies due to different acoustical changes (either measured or artificially added). In most of these papers, the authors proposed different approaches to design a system that can achieve a “reasonable” performance despite these inaccuracies, usually by the addition of more or less complex ways to regularise the cost function (at the expense of a performance deterioration [94]). However, one could also try to find a way for the system to always perform at its optimal regardless of the acoustic condition.

Another type of solution proposed in the literature is to incorporate information from the acoustic modifications to re-calculate or adjust the filters.

- **Measured temperature changes:** In 2017, Olsen *et al.* [26] studied how the changes in temperature affect the performance of the PSZ system in a real car. The authors proposed two compensations strategies, one to overcome the modification in the loudspeaker’s response and one to compensate the modification in the speed of sound.

In spite of the fact that the compensation strategies were incapable of achieving the performance when the acoustic system was accurately estimated, the authors proposed an original approach towards the

robustness of the PSZ. The idea behind was to incorporate physical information (in this case the temperature) to improve the accuracy of the IRs used for the filter calculation.

However, as in the system under study (Sec. 3.4) the transducers are not in contact with the outer panels (*e.g.* doors, ceiling, etc.) and due to the fact that nowadays it is a standard to have automatic climate control inside cars, it is assumed from this point forward that changes in temperature as abrupt as in Ref. [26] are unlikely to strongly modify the acoustic environment under study.

- **Listener positioning:** In 2016, Simón Galvez *et al.* [153] proposed to adapt the filters of a PSZ system as a function of the position of the zones, relative to the sources. Moreover, according to the listener's position -given by a listener tracking device- the authors proposed to apply a set of filters to modify the directivity of the array, these correction-filters were calculated analytically assuming that each loudspeaker behaves like a point-source monopole. Moreover, a similar approach was published by Ma *et al.* in 2018, incorporating crosstalk-cancellation [103].

A very innovative study was presented by Møller *et al.* in 2020 [13]. In this paper the authors introduced the use of the moving horizon framework to generate moving sound zones that follow the listeners in a room. This is not a deterministic approach as presented in this thesis. Nevertheless, this control framework enables adaptation to time-varying changes by predicting future states of the system given current control actions. The method relies on knowledge of the room impulse responses at the positions where the zones should move to, but, instead of pre-calculating and storing all the possible combinations, the moving horizon framework presents an on-line method to calculate the filters.

Although the two approaches described above resulted in very promising results, their applications rely on an accurate knowledge of the acoustic conditions (either to implement the moving horizon or to calculate the correction) which is not always available. Due to the variations between different cars and different car models, it would be necessary to have a correction for each individual car, for each passenger configuration, every seat/headrest combination, etc.

Another solution that has not been so widely reviewed *per se*, is the possibility to re-characterise the system which would lead to the “optimal” condition of the PSZ. More precisely, the capability of re-measure all the necessary IRs to re-calculate the optimal set of filters.

- **Noisy IRs:** A notable exception to the usual study on inaccuracies in the IRs/TFs is the one presented by Møller *et al.* in 2018 [128]. In this study, rather than aiming towards the design of one robust set of filters that will work regardless of the acoustic conditions, the authors focused on the fact that the TFs can be acquired *in-situ*, overcoming many of the inaccuracies mentioned above. However, it might not be possible to control the noise when these *in-situ* TFs are acquired. Therefore, they propose to use an effort regularisation based on the estimated noise in the TFs.

Despite that the addition of this type regularisation term is ignored in this thesis, it should be considered in future works.

## Discussion

A series of studies were reviewed considering different types of acoustic modifications. Among these studies, it was possible to define which modifications are most-likely to affect the performance of the application studied in this thesis.

On the one hand, the variations in the FRF of the loudspeakers might be possible to overcome with performing one *in-situ* measurement if one ignores the loudspeaker ageing and assumes that the loudspeakers are working in their linear regime. On the other hand, uncertainties in the source/zones positioning have been studied from a manufacturing point of view (product variability), however a very plausible acoustic change in a car cabin can be a modification in the seat or headrest position. In addition, it is also very-likely that there is a modification in the passengers in the rear seats.

Lastly, it was showed that several authors focused on the calculation of a universal set of filters at the expense of a reduction in performance. Other authors propose a correction of the filters as a function of the changes in the listener positions, however they usually rely on knowing the characteristics of the environment.

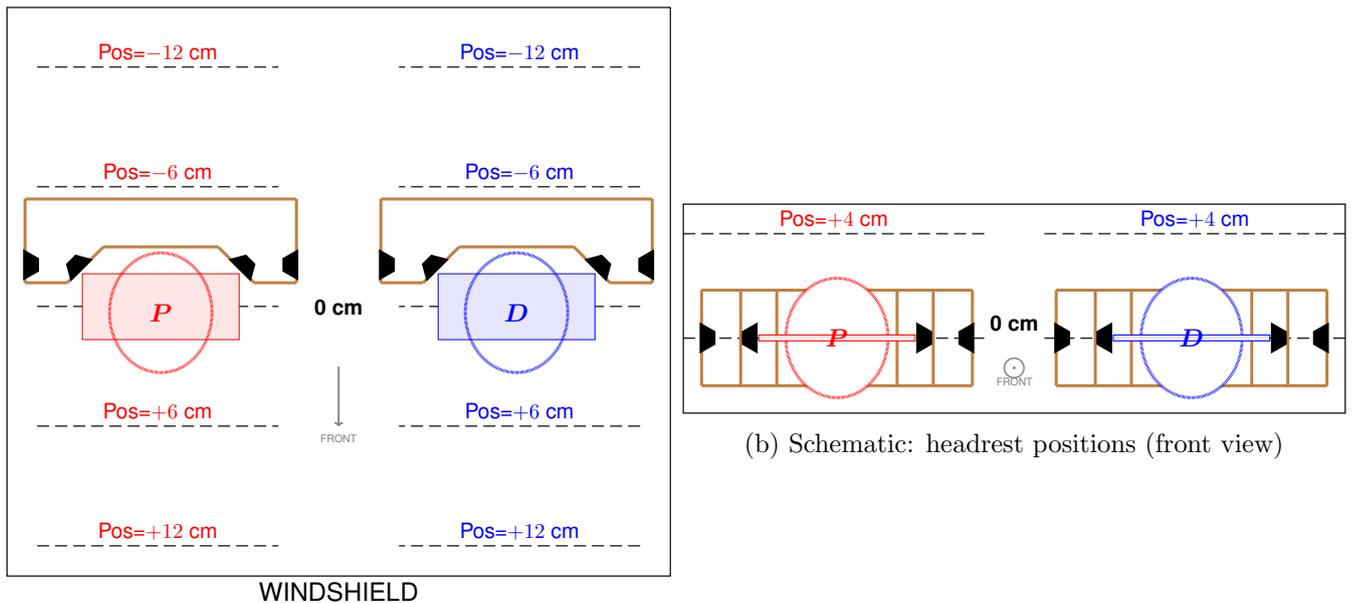
From this point forward, the focus is placed in trying to achieve the best performance the system can offer under the different acoustic conditions when the access to a very reliable model of the acoustic conditions is not available.

## 4.2 A library of impulse responses

In order to evaluate the aforementioned changes, one proceeded to build “library” of impulse responses using the prototype in the car cabin (Fig. 3.32) and changing the seat positioning, the headrest positioning and adding some mannequins to the back seats.

Therefore, this library of IRs was acquired considering the following modification in the acoustic environment:

- Five different seat positions:  $\pm 12$  cm with a step of 6 cm (see Fig. 4.1a).
- Two headrest positions: 0 cm and +4 cm (see Fig. 4.1b).



(a) Schematic: seat positions (top view)

Figure 4.1: Changes in the seat positions (a) and changes in the headrest elevation (b).

It is necessary to mention that the microphones are attached to the headrest, *i.e.* the zones to control move as the headrest moves. Therefore, when the seat/headrest position changes two things are modified, the relative positioning of the zones and the relative position of the zones respect to the cabin. In addition, when the headrest is elevated, the dummy is also elevated proportionally so that the loudspeakers are always at ears-height meaning that the relative positioning between the listeners and the loudspeakers does not change [23].

- Passenger configuration: following the works of Cho *et al.* [152], the effect of the passengers in the rear seat was evaluated adding two dressed mannequins in the back-seats (as in Fig. 4.2).



Figure 4.2: Dummies added in the rear seat

Moreover, for the passenger configuration only two configurations were analysed, with and without two passengers in the rear seats.

Furthermore, in order to avoid overestimation two sets of IRs were acquired, one to calculate the filter combination and one to evaluate its performance. Adding all the considered different combinations resulted in a library of more than 50 000 impulse responses.

### Calibration and evaluation positions

The aim of this chapter is to evaluate the effects on the PSZ system when there is a change in the seat and headrest positioning in the car cabin. To assess this study, two concepts derived from the practical application are introduced:

- Calibration position (CP): Seat and headrest positioning used as a reference to obtain all the IR's and calculate the optimal filters.
- Evaluation position (EP): Seat and headrest positioning used to evaluate the performance.

Moreover, at the end of the production line, it might be possible for the manufacturer to install a dummy with all the necessary microphones (as done in Chapter 2), perform all the necessary IR measurements and calculate the filters at a finite number of calibration positions or CPs. Once the car is in the street, it is necessary to evaluate the performance if the seat moves from the CPs, in other words, what happens in the evaluation positions (EPs).

## 4.3 Sensitivity evaluation

In Chapter 3, the system was studied under one specific acoustic condition, *i.e.* the seats in the centre of the cabin, headrest in the bottom position and no passengers in the back seats. However, it was shown in Chapter 2, that the resulting filters depend on the input characteristics. Furthermore, if one assumes that it is possible to re-measure the IR and re-calculate the optimal filters for every single acoustic condition, **Is the system still capable of fulfilling the constraints in all positions?**

Specifically, if one could have as many CPs as EPs, is the system capable of achieving comparable levels of  $AC$  and  $E_{rr}$ ? From this point forward this situation is referred to as the “Re-calibrated” approach (what Nielsen *et al.* refer to as having “oracle-knowledge” [134]).

Secondly, it might not be practical -or necessary- to have as many CPs as EPs, so one could wonder if system's performance is affected by the aforementioned changes, in other words, **if the seats or the headrest move from the CP, is the performance lost at the new EP?**

To answer the second question, it is assumed that only one CP is possible and maintained for all the EPs, from this point forward referred to as the “Fixed” approach. The CP is defined as the position studied in Chapter 2, when both seats are in the centre of their rails and the headrests are in the bottom position.

For this evaluation, assuming there are no calculation time constraints, the time domain formulation from Sec. 2.3.3.3 is used with a filter’s length of 200 coefficients.

### 4.3.1 Seat position change

Firstly, one proceeded to evaluate only the effect of the seat positioning. From the library generated in (Sec. 4.2), as both seats can move independently (see Fig. 4.1a), it is possible to obtain 25 different seat combinations.

To simplify the analysis, in this report it is analysed when both seats move symmetrically *e.g.* if the driver seat moves to  $-12$  cm, also the passenger seat moves to  $-12$  cm. These movements were chosen because they represent among the worst deteriorations, but they are not the only ones studied, other seat-changes can be found in App. E, for example when only the passenger seat moved, or when only the driver seat moved, etc.

Firstly, one proceeded to study the resulting  $AC$  and  $E_{rr}$  for the five available positions taking as an example, the CP in the centre position.

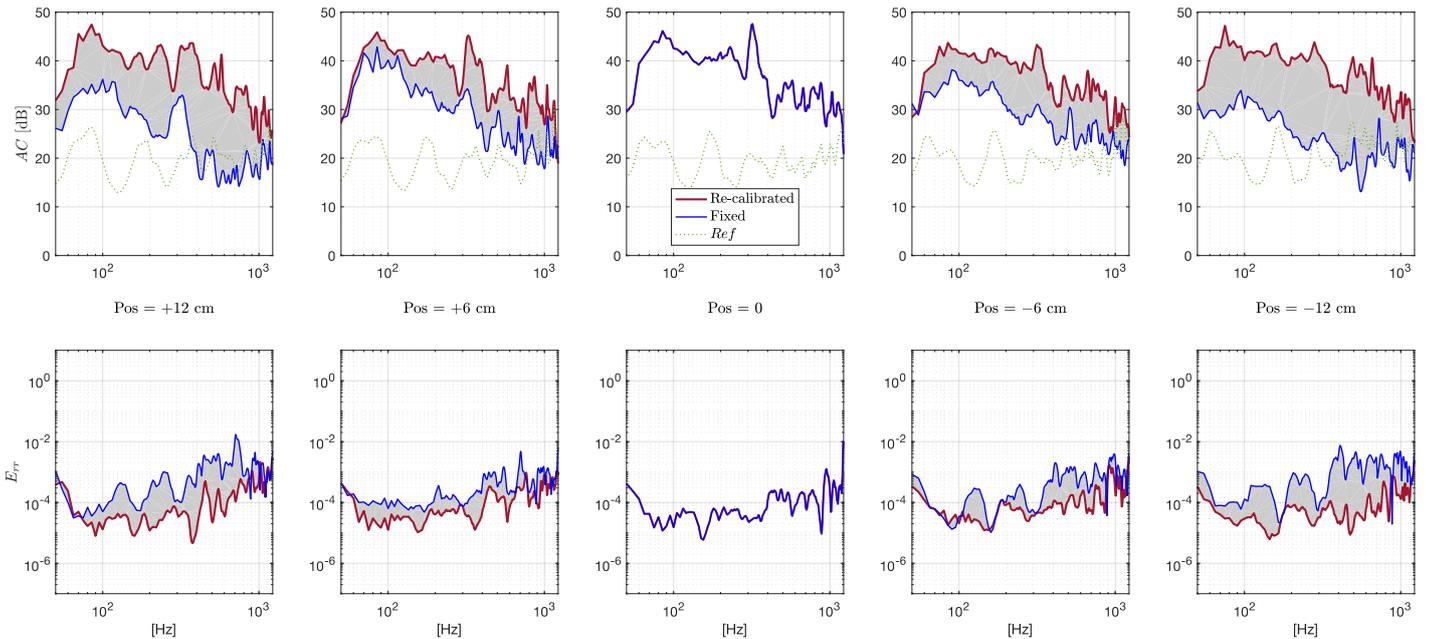


Figure 4.3: Resulting  $AC$  and  $E_{rr}$  for symmetric seat changes. Each column corresponds to one seat position (see Fig. 4.1a). Both the re-calibrated approach as well as the Fixed using as CP the centre position (Pos = 0) are showed.

Both questions posed at the beginning of the chapter can be answered by analysing Fig. 4.3. Firstly, the re-calibrated approach gives comparable performance regardless of the position of the seats, *i.e.*, the system is capable of fulfilling the constraints in all five positions. It is possible to observe that the system is achieving levels of  $AC$  of the order of 30 dB or more for the bandwidth of interest at all positions. Additionally, the resulting  $E_{rr}$  in all positions maintain levels of the order of  $\simeq 10^{-4}$ .

On the other hand, one can observe the deterioration of the “fixed” approach when the seats move from the centre position. In other words, the filters calculated at CP are no longer valid when the seats move, especially at positions  $\pm 12$  cm above 400 Hz, the CF result in even worse  $AC$  than without applying any processing (worse than the *Ref*).

To sum up, the system is capable of maintaining the performance in all the positions but it needs a dedicated filter combination per seat positioning.

Both the “re-calibrated” as well as the “fixed” approaches define a corridor (shadowed area between the curves in Fig. 4.3). The “re-calibrated” approach defines the ceiling of the performance, the best performance achievable for this system (for this cost function, this formulation, this environment, etc.).

The “fixed” approach defines the floor of performance, which does not refer to the worst performance achievable. When designing a system capable of changing/adapting the filters, the fixed approach establishes the performance that this new system needs to outperform, otherwise there would not be any improvement.

It is necessary to mention that the studied modification does not modify the relative positioning between the zones and/or the sources (as their rails are parallel, their relative position does not change). However, there is a modification in the relative position between the system (loudspeakers + zones) and the cabin itself, *i.e.* a modification in the way the modes are driven, a modification in the way the modal response is being received, etc.

As mentioned in Sec. 3.4, one could try to explain the main peaks and depths of the resulting metrics, but this analysis has become even harder than before. For example, in Fig. 4.4 one shows a superposition of the different *Refs* at each seat position.

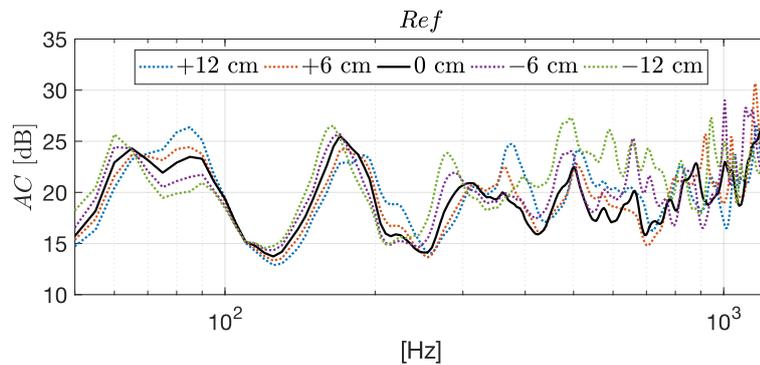


Figure 4.4: Resulting *RAC*, for symmetric seat movements in a car cabin.

From Fig. 4.4, it might be possible to attempt a modal analysis up to  $\approx 250$  Hz for the *RAC* (see Fig.4.4). However, the car cabin is not a “shoe-box” (its width and height are reduced towards the front), but the fact that a large obstacle is moving along the cabin (the seat) changes the frequencies and the modes shapes of the cabin [154, 155]. Hence, it would be necessary to have the modal response per position of the seat. In addition, one needs to study where the *DZ* is placed (if it is in a maximum of pressure, in a nodal line, etc.) and, as the *BZ* is also moving, it is also necessary to analyse how the cabin’s modal response is being driven.

Lastly, the changes in the *RAC* (Fig. 4.4) are considerably smaller than the deterioration in performance observed when the filters are applied (Fig. 4.3).

### Bonus: Seat extremes

As a bonus, one proceeded to also analyse the two most extreme case in the seat positioning (see Fig. 4.5).

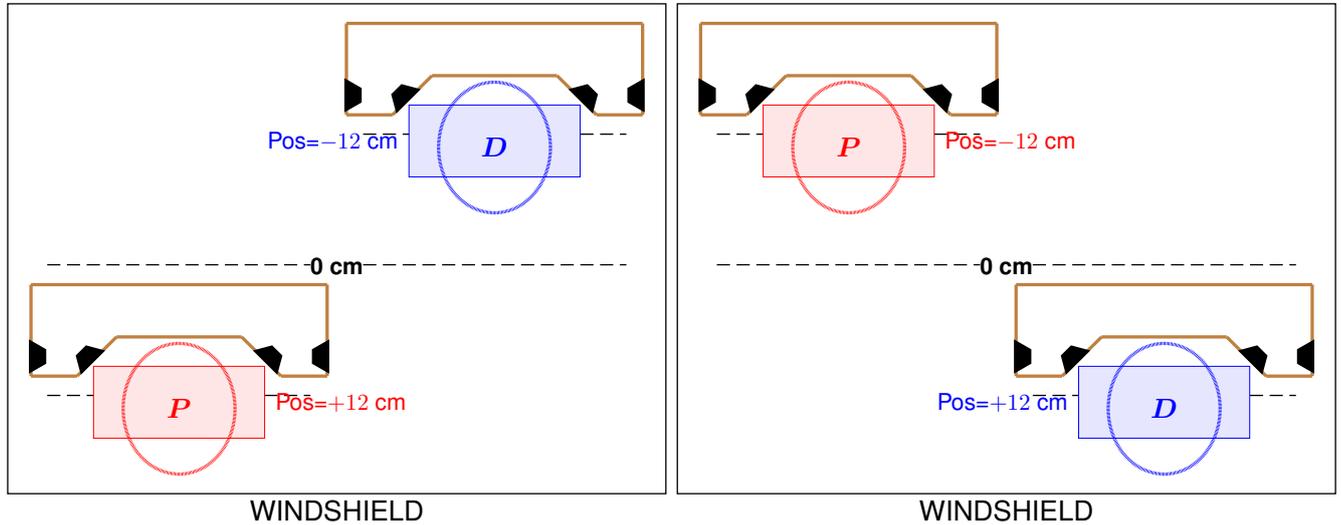


Figure 4.5: Schematic: seat positions extreme positions.

These seat-positioning are reviewed because it allows to affect the positioning of the seat regarding the cabin (drive differently its modal response) but it also exemplifies changes in the relative positioning of the sources and zones, not present in the symmetric-movements case.

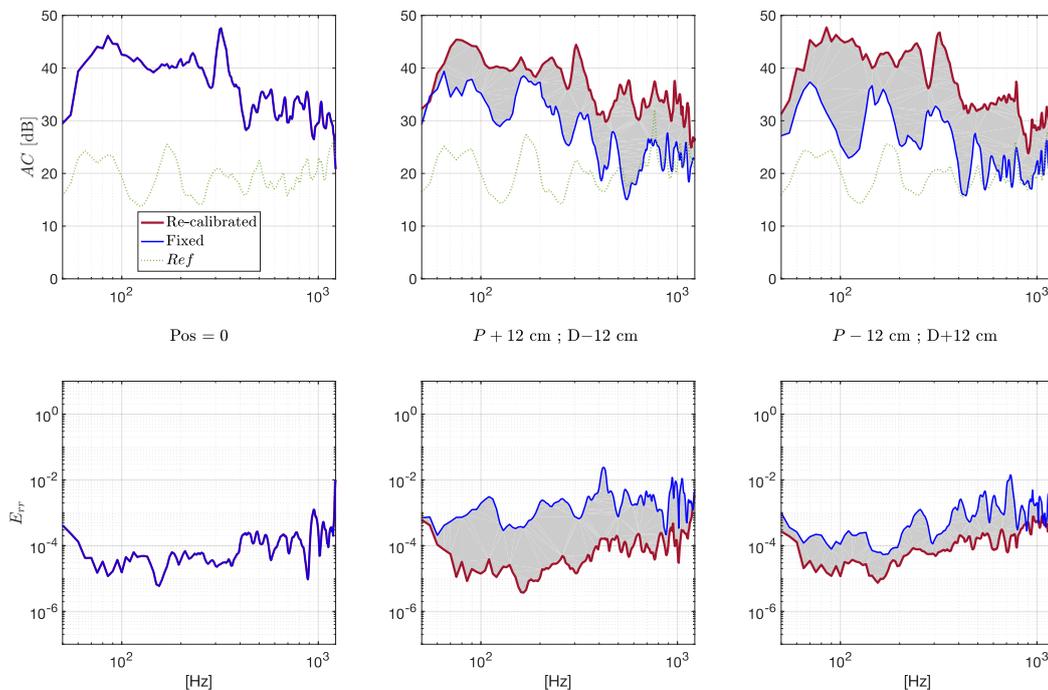


Figure 4.6: Resulting  $AC$  and  $E_{rr}$  for extreme seat positions. Each column corresponds to one seat position (see Fig. 4.5). Both the re-calibrated approach as well as the Fixed using as CP the centre position ( $Pos = 0$ ) are showed.

Similar conclusions as before can be drawn from Fig. 4.6. The re-calibrated approach is still capable of maintaining a relatively constant performance and the fixed approach adds evidence to the fact that a dedicated set of filters is needed when the seat positions change. Nevertheless, when comparing this results with the one in Fig. 4.4 one can observe similar  $AC$  deterioration but a larger deterioration in the  $E_{rr}$  when the passenger seat (the  $BZ$ ) is close to the wind-shield.

### 4.3.2 Headrest position change

In addition to the seat positioning, one proceeded to study the influence on the  $AC$  and  $E_{rr}$  when the passengers modify the headrest positioning. To this end, the CP was assumed to be when both headrests are in their lowest position (at 0 cm in Fig. 4.1b) and the complete study was made with the seats placed in the centre of the seat-rails (as in Sec. 3.4).

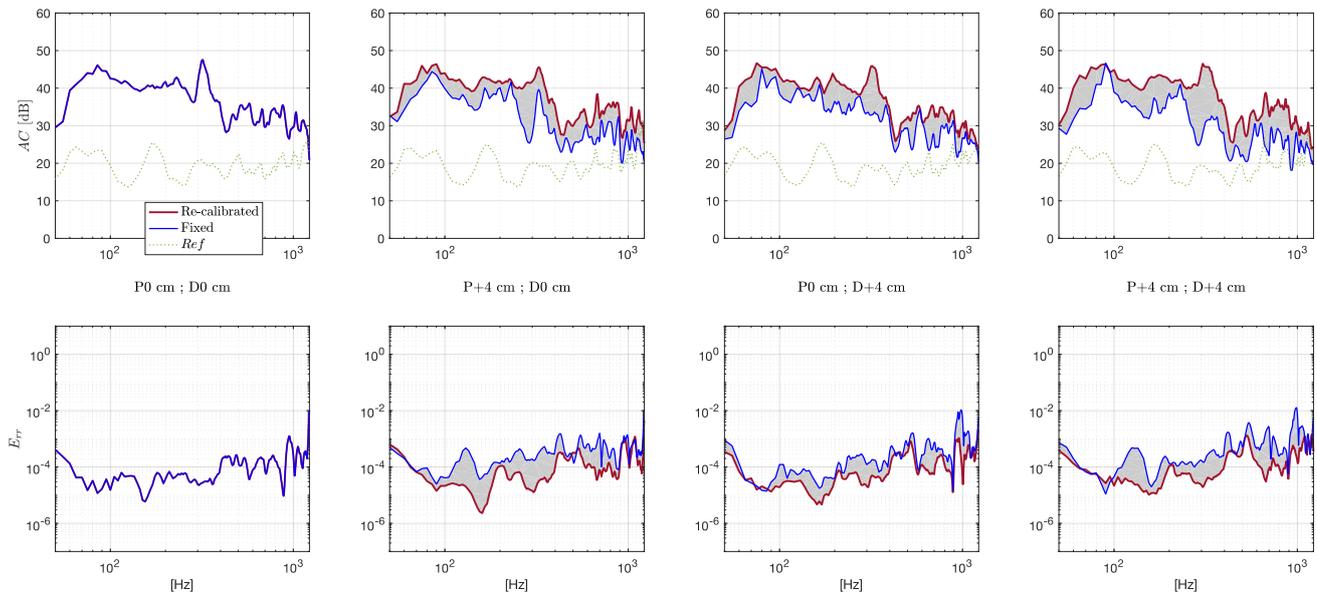


Figure 4.7: Resulting  $AC$  and  $E_{rr}$  for changes in the headrest positions with the seats at ( $Pos = 0$ ). Each column corresponds to one headrest positioning (see Fig. 4.1b). Both the re-calibrated approach as well as the Fixed are showed using as CP the centre position ( $Pos = 0$ ) when both headrest are in the lowest position.

As in the seat-positioning, the system is capable of maintaining its performance for all four configurations using the Re-calibrated approach. Additionally, a deterioration of both the  $AC$  and the  $E_{rr}$  when only using the CP filters -Fixed approach- is also present.

Nevertheless, it is possible to observe that the seat positioning generates a much worse deterioration in the overall-performance of the system (Fig. 4.4). It is interesting to observe that at some particular frequencies, the performance is almost unaltered (230 Hz at P+4 cm - D0 cm, 80 Hz at P0 cm - D+4 cm, 90 Hz at P+4 cm - D+4 cm, and others), *i.e.* there are not acoustical changes at these particular frequencies for these particular positions.

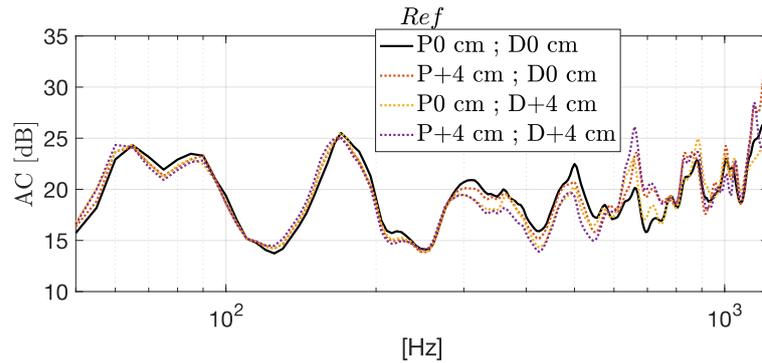


Figure 4.8: Resulting  $RACs$ , for different headrest positioning in a car cabin.

When overlapping all the  $RACs$  (Fig. 4.8), although it is possible to observe that there is a very small variation between the different headrest positions there is still the need of adapting the filters when the headrest moves to maintain the re-calibrated performance (Fig. 4.7).

### 4.3.3 Passenger configurations

Lastly, one proceeded to evaluate the sensitivity of the system when two passengers are present on the rear seat.

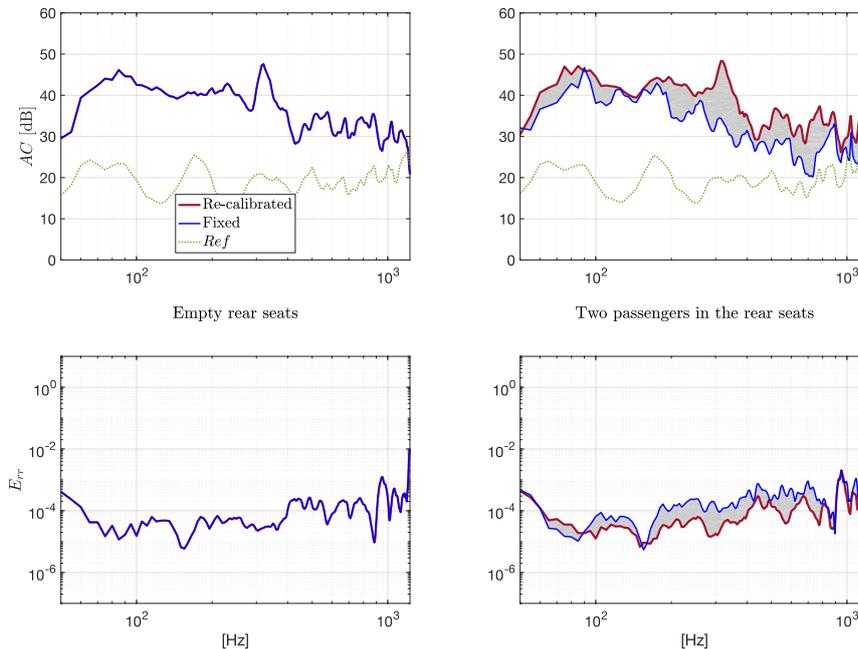


Figure 4.9: Resulting  $AC$  and  $E_{rr}$  with and without passengers in the rear seats. Both the re-calibrated approach as well as the Fixed (filters calculated without passengers) are showed.

Once more, it is possible to observe that adding two adult-size mannequins in the rear seats also deteriorates the performance of the system. The passenger configuration modification is another example of a modification in the acoustic environment without a modification of the relative positioning between the zones and sources.

It is important to notice that, that the strongest modification can be seen in the  $AC$  whilst the  $E_{rr}$  deterioration is very small.

### 4.3.4 Summary

In this section, a study on the sensitivity of system -designed in Chapter 3- to different acoustic conditions was performed. On the one hand, the good news is that the system designed under static conditions is capable of achieving a relatively constant performance regardless of the acoustic conditions -different seats positions, different headrest positions and extra passengers-. The near-field array is still capable of achieving levels of  $AC$  of the order of 30 dB for the band of interest.

On the other hand, the bad news is that unless a suited filter combination is calculated for each acoustic condition, there is a deterioration of the overall performance of the system, especially the  $AC$ , in some cases even worse than not applying any processing at all.

Regarding the  $E_{rr}$ , it is possible to observe that the deterioration is considerably smaller. This can be attributed not only to the near-field location of the sources, but also to the fact that the way the problem is posed, the reproduction filters do not change (this is re-addressed in detail later).

In spite of its obviousness, in order to achieve large performances, the accuracy of the IRs is a key factor especially for PSZ which rely on destructive interference [128, 13]. Hence, from this point forward the question to answer is not how to find a filter combination that will maintain an acceptable performance for different acoustic changes, but **how to obtain the optimal set of filters for the different acoustic conditions?** to achieve the best-performance the system can offer.

In addition, it has been highlighted in this section the “nature” of acoustic modification, *i.e.* if the modification either changes the relative positioning between zones/loudspeakers, if it is solely a modification of the acoustic environment or both things simultaneously. It can be stated from the different modifications reviewed that regardless the modification’s nature, it is necessary to assign a dedicated set of filters to maintain the performance.

Lastly, it is important to mention that all the modifications reviewed in this thesis share the common feature that, after a modification occurs, the system can be seen as a stationary environment, unlike the moving sound zones reviewed for example in Ref. [13].

The following studies are made using the five symmetric seat positions (from Sec. 4.3.1) to evaluate different strategies because they offer a wide range of performance deterioration (Fig. 4.3). Nevertheless the following study can be extrapolated to -ideally- any acoustic modification.

Firstly, one is going to review two direct applications of the deterministic approaches:

- Database solution: if one had in storage most of the possible filter combinations it would be possible to simply choose which filter to apply to each configuration (Sec. 4.4).
- Re-measurement solution: Due to the growing availability of voice-controlled technologies and built-in hands-free technologies in cars, one could find the way to re-measure the IR and have a computer capable of re calculating a new filter-combination when there is a loss in performance for example (Sec. 4.5).

The advantages, disadvantages, complications of these two solutions are analysed in the next sections as well as some contributions to make these solutions more attractive for the car-application.

## 4.4 Database solution

More and more the car’s computer is aware of everything that is happening in the car cabin, not only to know in real time the temperature of the cabin (automatic thermostats), but also the position of the seats (electrically controlled seats), the number of passengers (weight sensors and cameras), etc. Therefore, if one had in storage all the possible filter combinations needed, one could simply choose which combination to use and apply it as a feed-forward control technique [3]. Hence, the most “straight-forward” solution would be to implement a database solution.

This type of solution can not be directly found in the literature, not because it has not been studied, but because it is a direct application of the deterministic approaches reviewed in Sec. 2.3.1. Moreover, looking

at the block diagram in Fig 2.13, this solution can be understood as anticipating every input combination that can happen. This is found in the literature for example as Nielsen’s “*oracle-knowledge*” term [134] or Møller’s “*create multiple time-invariant solutions*”[13].

Furthermore, one could assume that all the necessary calculations (with the corresponding measurements) are performed at the end of the production line with a dedicated microphone-array and external processing. Therefore, there is no need to have any dedicated microphones installed in the car cabin or extra computational resources to be able to calculate the filters “on the go”.

However, due to the variations in the manufacturing process as well as the sensitivity of the methods due to inaccuracies in the TFs/IRs (reviewed in Sec. 4.1 and analysed in Sec. 4.3) it is necessary to calibrate each car at the end of the production line.

Additionally, it would be necessary to run a calibration-measurement for all the possible combinations of the acoustic environment and calculate the corresponding filters, which would take an extremely long calibration time per-car at the end of the production line. Furthermore, assuming that all the variations could be anticipated, it would be necessary to incorporate a storage unit in the car cabin capable of saving all the filter combinations.

#### 4.4.1 Interpolation: saving time and storage

One way to help the implementation of the database solution is to be able to reduce the number of measurements needed at the end of the production line. Assuming to know three CP, two extremes (Pos  $\pm 12$  cm) and the centre (Pos = 0), one could interpolate the TFs from the “unknown” positions ( $\pm 6$  cm).

To test this hypothesis, one proceeded to interpolate the TFs using a linear interpolation and a quadratic interpolation. Moreover, for each loudspeaker/microphone TF, the real and imaginary parts were interpolated separately and then added.

Therefore, the resulting  $AC$  and  $E_{rr}$  are studied in all five seat-symmetric positioning.

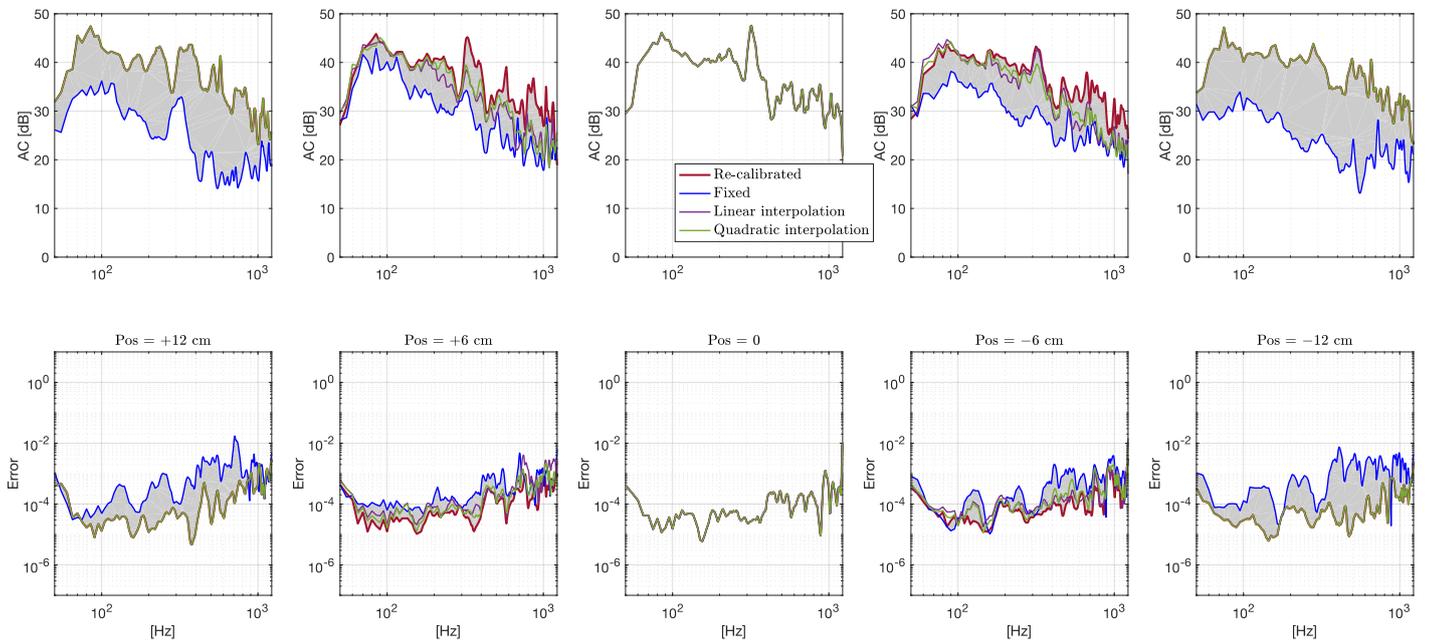


Figure 4.10: Resulting  $AC$  and  $E_{rr}$  for symmetric seat positions. Each column corresponds to one seat position (see Fig. 4.1a). Three CP are taken at  $+12$  cm,  $0$  cm,  $-12$  cm, and the filters in middle positions are calculated using interpolated TFs.

From Fig. 4.10, one can observe that at  $\pm 6$  cm, the interpolated filters achieve an improvement of the  $AC$  up-to  $\approx 350$  Hz where the response starts falling up to  $900$  Hz -  $1$  KHz where the responses match

the Fixed approach. Regarding the  $E_{rr}$ , it is possible to see a similar behaviour, the interpolated filters present a better response than the Fixed approach up to  $\approx 350$  Hz where it loses accuracy.

#### 4.4.2 Database conclusions

Assuming one has the best methodology to calculate the filters, the database solution presents without any doubt the best performance the system can achieve. However, it presents a series of complications that can make it not-viable for the car application.

Firstly, it is necessary to anticipate all the possible acoustic conditions that might happen in a car and find the time to obtain all the necessary filter-combinations (acquisition + filter calculation).

As an attempt to make this solution viable, one proposed a simple interpolation between seat-positions in order to calculate the system's IRs at un-measured EP. This trick was able to maintain the system results up to  $\approx 350$  Hz where the modal density of the car cabin is relatively low (typically up to 300 – 440 Hz [61, 62]) and  $\lambda$  is large compared to the seat displacement.

Furthermore, in the case of the seat-positioning the “mid” positions for interpolating the IRs are simpler to find, however it is necessary to also establish all the possible combinations between the different modification (headrest position, passengers, etc.) and define a grid of “middle-points” that can be interpolated.

It has been proposed in the literature to use the common-acoustical-pole and residue model to spatially interpolate and extrapolate room TFs [156]. However, these methods propose that the acoustical poles are invariant in a given room (which can be restrictive in a car) and it has been successfully applied mainly for the lower frequencies in rectangular rooms. The study of the implementation of the common-acoustical-pole and residue model is out of the scope of this thesis.

Additionally, due to all the possible changes that happen in a car cabin, one could wonder if the time and storage saved by a simple interpolation is enough (regardless of the performance lost).

Lastly, this solution presents the following advantages and disadvantages:

##### Advantages:

- **Simplicity:** It is just a direct application of the deterministic approaches.
- **No microphones in the cabin:** The calibration is performed in the factory, no need of dedicated microphones.
- **No re-calculation:** There is no need to re-calculate the filters, it is just necessary to choose among multiple time-invariant filters [13].

##### Disadvantages:

- **Long calibration:** At the end of the production line, it is necessary to measure and calculate all the possible acoustic changes that might happen in a car cabin.
- **Storage:** It is necessary to have a dedicated storage unit to store all the filter combinations.
- **Anticipate:** It might not be possible to anticipate every acoustic modification. Some might be not so simple to anticipate *e.g.* loudspeaker wear.

## 4.5 Re-measurement solution

Another direct application of the deterministic approach is the “re-measurement” solution. When there is a change in the acoustic conditions that results in performance-losses (either  $AC$  or  $E_{rr}$ ), the system can run a re-acquisition of the IRs and re-calculate the new filter combinations according to the new conditions.

The microphone outline used up to this point was chosen in such a way that it was capable of sampling the complete problem, however, it is very costly to have 36 microphones in a car (transducers, cables, converters, etc.) and, when looking at their outline (Fig. 4.11), it is not realistic to assume that one can install microphones so close to the head.

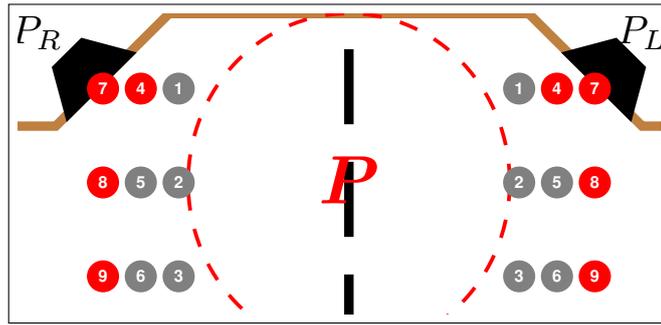


Figure 4.11: Schematic: microphone outlines passenger seat.

Therefore, the main challenge to apply the re-measurement solution is to reduce the number of microphones according to their positioning.

As it is necessary to re-calculate the filter combination every time there is a loss of performance, the calculation time taken by the algorithm to obtain the filters is a key parameter in this solution, hence, filters are calculated using the frequency domain formulation (see Eq. 2.33). For simplicity, one is only going to consider head-symmetric outlines (as in Fig. 4.11) and an identical outline is used in the other seat, *i.e.*, when one refers to for example microphone 7, it means the “four microphones 7” in both seats.

#### 4.5.1 Microphone-reduction

If one observes Fig. 4.11, microphones 1, 2, 3, 5, and 6 (in grey) are positions that in practice, for safety reasons, are not possible to achieve physically. On the other hand, microphones 4, 7, 8 and 9 (in red) might be possible to install in the headrest or to achieve a similar position in the cabin. Therefore, one proceeded to test the performance of the system when only using microphones 4, 7, 8 and 9 (red microphones in Fig. 4.11 and identical for the other seat) and all the possible combinations aiming towards the reduction of microphone channels.

Firstly, it is necessary to review the limitations/repercussions of the microphone reduction:

- Numerical limitation: in order to calculate the filters using the PM it is necessary to have an over-determined system (see Sec. 2.3.3.3), *i.e.* the number of microphones needs to be larger than the number of sources ( $L \leq K$ ). Therefore, maintaining the symmetry of the system, one proceeded to study the microphone outlines listed in Table 4.1.

Id	$K$	Mic. outline
a)	36	1 to 9
b)	16	4, 7, 8 and 9
c)	12	4, 7 and 8
d)	12	4, 7 and 9
e)	12	4, 8 and 9
f)	12	7, 8 and 9

Table 4.1: Microphone outlines for removal

Additionally, to avoid any ill-conditioning problems, a frequency-dependent regularisation is added (see Sec. 2.3.3.4) in such a way that the Cond. is limited to a maximum of 3000.

- Physical repercussions: The fact that control microphones are being removed, means that the assumption that the IR’s fully/perfectly parametrize the zones (see Sec. 2.3.1) partially-losses its validity; *i.e.* one is no longer controlling the complete zone, but only controlling at some particular locations.

To improve visibility, it is possible to assume that the re-calibrated approach achieves the best results (for this application), hence one can calculate the  $AC_G$  of the different outlines presented in Table 4.1 taking as a reference outline  $a$ ).

$$AC_G|_i(\omega) = [AC|_i(\omega) - AC|_{K=36}(\omega)] \quad (4.1)$$

being  $i) = a), \dots, f)$  the different outlines considered.

Therefore, following the conclusions from Sec. 4.3, the new microphone configurations (listed in Table 4.1) are reviewed as a function of their resulting  $E_{rr}$ , Cond. and their  $AC_G$

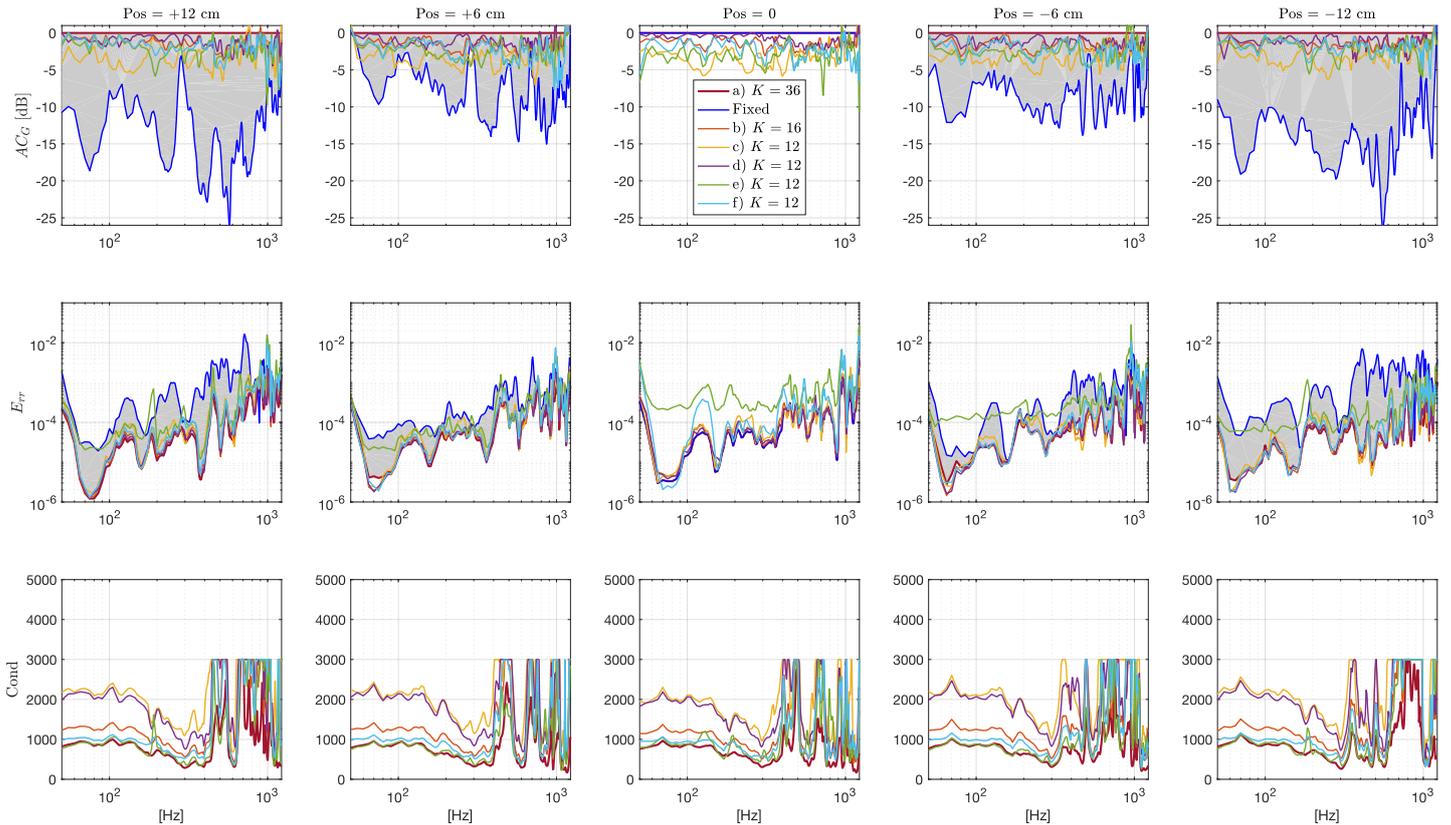


Figure 4.12: Resulting  $AC_G$ ,  $E_{rr}$  and condition number for symmetric seat positions. Results employing the re-measurement solution, using different microphone outlines.

When analysing the  $AC_G$  in Fig. 4.12, it is possible to conclude that it is always better to re-measure the IR -regardless of the microphone outline used- than just applying the filters from the calibrated position (Fixed).

Among the  $K = 12$  configurations, it is possible to see that the best -overall- performance is achieved by the  $d$ ) microphone-outline (Purple line in Fig. 4.12). Additionally, this outline achieves even better levels than  $b$ ) that has more microphones. This is an indicator of the aforementioned physical repercussions; due to the fact that one is no-longer sampling the complete zone (as is in outline  $a$ ), it is more important to choose the good location of the microphones than adding more microphones. Furthermore, it is possible to observe that the smallest  $AC_G$  results with  $c$ ) which happens to be the only outline without microphone 9 (see Table 4.1).

Regarding the  $E_{rr}$  it is possible to observe that all the microphone outlines outperform the Fixed except for  $e$ ) at Pos=  $-6$  cm, Pos = 0 and for some frequencies at the other positions. The only difference that this configuration has with the others is that it does not take into consideration the microphones the closest to the loudspeakers, microphone 7 (see Table 4.1).

Although it is possible to find frequencies where the  $AC_G > 0$  (e.g. at Pos = 0  $\approx$  945 Hz), this does not mean that this system is better, it means that it is easier for the algorithm to obtain the potential energy cancellation than the error minimisation (at the same frequencies there is always a growth of the  $E_{rr}$ ).

Lastly, it is possible to observe that the algorithm has to impose the effort constraint in all the cases with  $K = 16$  and  $K = 12$  after  $\approx 400$  Hz where the condition number exceeded the limit value.

To sum up, in order to guarantee a small  $AC$  deterioration, the outline needs to have microphone 9 and to keep the  $E_{rr}$  as low as possible it is necessary to have the information from microphone 7 (see Fig. 4.11), which corresponds respectively to the microphones the furthest from the loudspeaker (more information about the acoustic environment) and the closest to the source (only direct field).

Moreover, in the literature it is found that it is necessary to have an over-determined system, but if all the TFs -for all the frequencies- are linearly independent of each other, it should be possible to have  $L = K$  [124]. Therefore, one proceeded to calculate the  $E_{rr}$  and the  $AC$  when only leaving microphones 7 and 9. To illustrate the effect of regularisation, the results both with and without regularisation are shown.

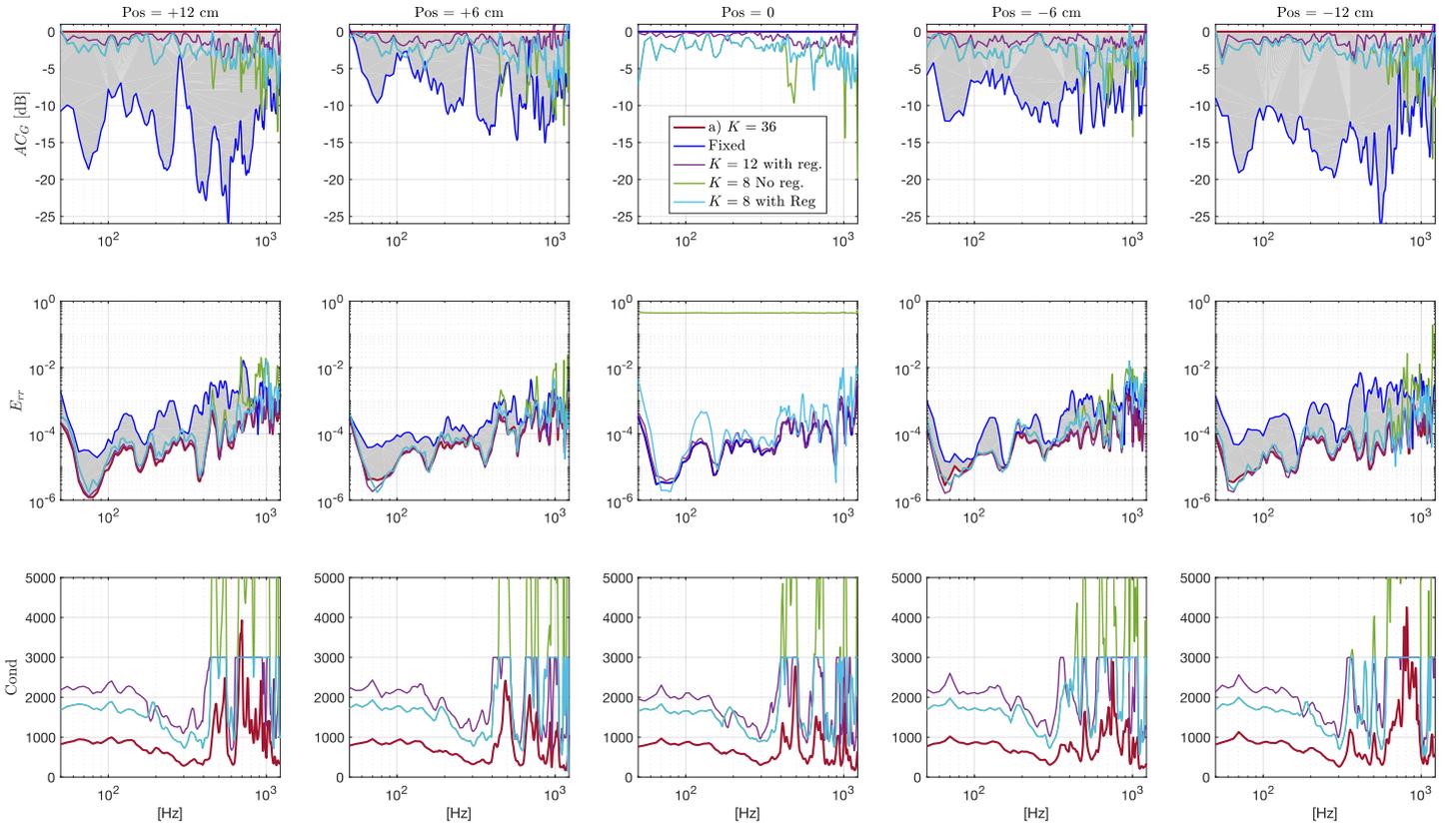


Figure 4.13: Results: re-measurement solution,  $K = 12$  vs.  $K = 8$  with and without regularisation.

In spite of the fact that  $K = 8$  is not performing as-well-as  $K = 12$ , it is possible to observe that it is still capable of improving the performance in all five positions (see Fig. 4.13). At Pos = 0, one can observe that the  $K = 8$  without regularisation is not being capable of calculating the solution (an error close to 1). In addition, the use of regularisation improves the performance at the frequencies where it is applied (above  $\approx 400$  Hz).

## 4.5.2 Microphone-reduction extension: Virtual Microphones

In the PSZ literature, one example to reduce the number of microphones used for the acquisition of the TF was introduced by Zhu *et al.* [150] where only one TF is acquired per loudspeaker and per zone and new TFs were extrapolated using the radiation model of a monopole source to more points sampling the zones and the mismatches were compensated by acoustic modelling based robust control. Although this study presented interesting results, with almost identical  $AC$  values, in order to apply this model-based extrapolation for the headrest application, due to the near-field location of the zones and the diffraction of the head and shoulders it would be necessary to incorporate a more complex model, this is out of the scope of this thesis.

Another option that has not been used in the PSZ literature (to the best of the author's knowledge), but that has been widely studied for ANC, is to use the principle of virtual sensing algorithms in order to increase the number of microphones by adding virtual/remote microphones to the calculation [157].

The main principle of these algorithms is to install a number of sensors in an appropriate position -the physical microphones ( $k_p$ )- and to have the optimisation happen at a different location where the sensors cannot be installed -virtual microphones ( $k_v$ )- [158]. It is possible to find numerous papers on the study of virtual sensing algorithms (many reviewed in Ref. [157]) and some specifically applied to loudspeaker-headrests, *e.g.* [159, 160, 161].

The algorithm proposed in this thesis is inspired by the remote-microphone technique introduced by Roure and Albarrazin in [158]. However, there are some very important differences between this method and the application proposed in this section. In the ANC developments, the virtual sensing algorithms are used to estimate the error signals (from the physical microphones to the virtual microphones) [157] whilst in this section, it is used to extend the IR/TF matrices. Additionally, in this study both physical and virtual microphones are used in the optimisation, rather than only virtual.

### Proposed method

If one has one physical microphone installed in the headrest  $k_p$  and one virtual microphone  $k_v$  in a position that can not be achieved physically, the transfer function between loudspeaker  $l$  and the aforementioned microphones can be written as  $G_{k_p,l}(\omega)$  and  $G_{k_v,l}(\omega)$  respectively. Assuming one can perform one calibration under specific acoustic conditions (one seat position, one passenger configuration, etc., *i.e.* at CP) where the "virtual" locations can be measured (*e.g.* with a removable microphone array) then it is possible to calculate the TF between the microphones for loudspeaker  $l$  as

$$V_{k_v|k_p,l}(\omega) = \frac{G_{k_v,l}^{CP}(\omega)}{G_{k_p,l}^{CP}(\omega)} \quad (4.2)$$

being  $V_{k_v|k_p,l}(\omega)$  the TF between  $k_v$  and  $k_p$  for loudspeaker  $l$  at the angular frequency ( $\omega$ ) measured at CP. Moreover, in theory Eq. 4.2 can be calculated for as many loudspeakers and physical / virtual microphones as needed and stored.

Afterwards, once the car out of the factory, it is assumed that it is possible to measure  $G_{k_p,l}^{EP}(\omega)$  for the new EP. Therefore, assuming that the microphone TF ( $V_{k_v|k_p,l}(\omega)$ ) remains constant for every acoustic modification in the car cabin ( $V_{k_v|k_p,l}(\omega)$  obtained at CP is constant for every EP), it is possible to retrieve  $G_{k_v,l}^{EP}(\omega)$  at EP from Eq. 4.2 as

$$G_{k_v,l}^{EP}(\omega) = V_{k_v|k_p,l}(\omega)G_{k_p,l}^{EP}(\omega). \quad (4.3)$$

In order for the assumption of  $V_{k_v|k_p,l}(\omega)$  to be valid regardless EP, it is necessary to choose the best microphone combinations as well as the best calibration position.

### Application

Therefore, one proceeded to exemplify the use of the microphone-virtualisation method proposed by adding one virtual microphone per ear, increasing from  $K = 8$  to  $K = 12$ . Moreover, the physical/virtual pairs are chosen such that the twp microphones are the closest. In this case to calculate virtual microphone  $k_v = 4$ , the physical microphone  $k_p = 7$  from the same ear is used (see Fig. 4.11). Furthermore, one CP at Pos = 0 is used, where  $V_{k_v|k_p,l}(\omega)$  are calculated for each  $k_v/k_p$  pair, for the eight sources and for the studied frequency band.

$K$ Physical	$K$ Total	$k_p$	$k_v$
8	12	7 and 9	4

Table 4.2: Microphones used for virtualisation.

One difficulty encountered when calculating the resulting virtual TF was that at some particular frequencies it was possible to observe some “spikes” due to the ratio of TFs in Eq. (4.2) (this is re-addressed below). In order to avoid these problems, one proposed to smooth the  $V_{k_v|k_p,l}(\omega)$  TFs -calculated at the CP- applying the moving median along the frequency band, one for the real part and one for the imaginary part.

All these aspects are illustrated in Fig. 4.14 by plotting the resulting TFs between loudspeaker  $P_R$  and the four virtual microphones (microphones 4) for five different symmetric seat-positioning (from Fig. 4.1a), displayed with a 25 dB shift between the positions to simplify the plot.

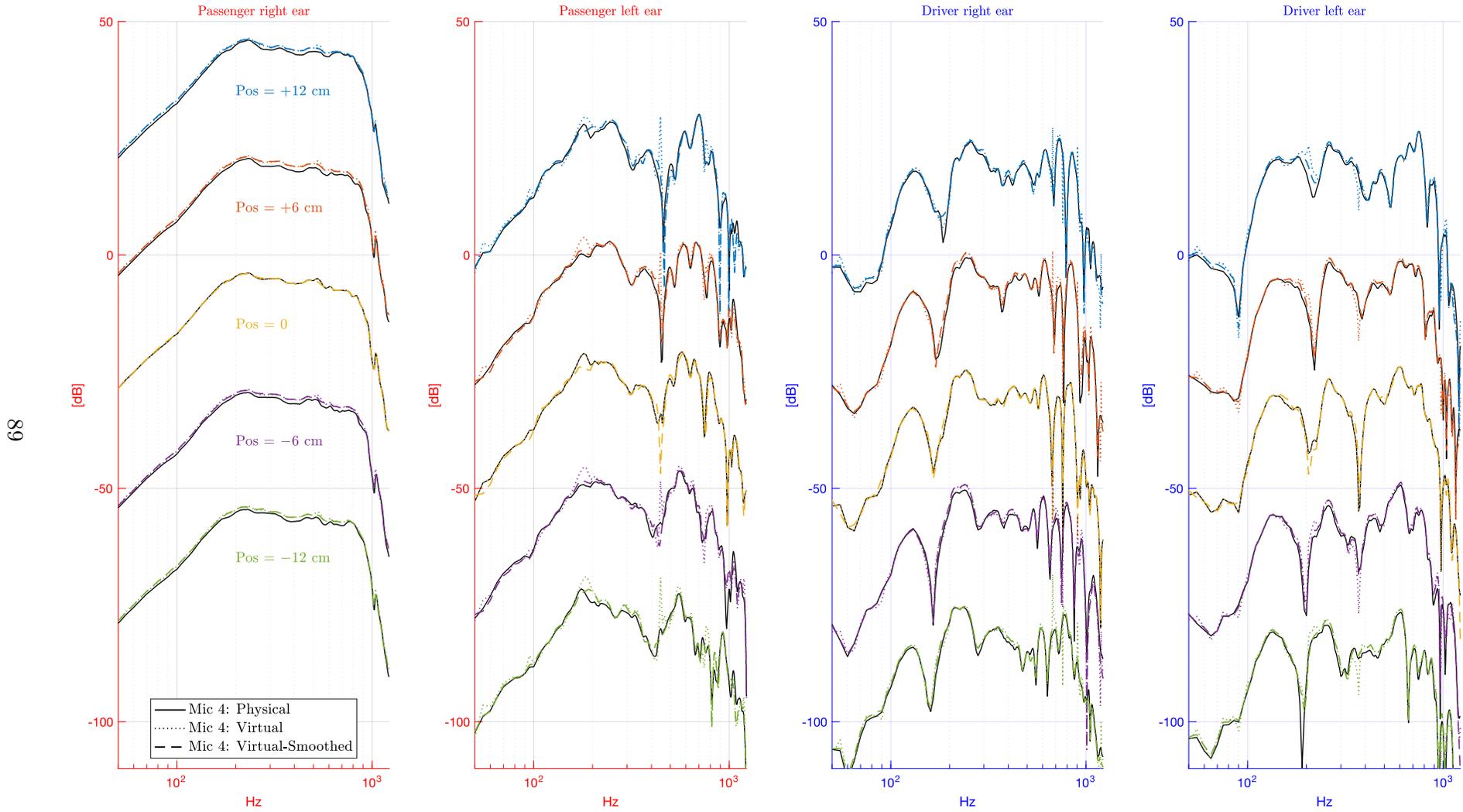


Figure 4.14: Results: Resulting amplitude response of the TFs between Loudspeaker  $P_R$  and the “virtualised” microphones 4 for five symmetric seat positions (with a **25 dB shift between positions**).

When analysing Fig. 4.14 it is possible to state that the virtualisation method is capable of approximating the TF even for microphones further from the loudspeaker where there is a stronger influence of the environment -modal response, diffraction, etc.- as it is the case of the microphones at Driver’s right and left ears. Even if no-smoothing is used, due to the proximity of microphones physical and virtual microphones (4 and 7 in Fig. 4.11), it is possible to state that  $V_{k_v|k_p,l}(\omega)$  stays almost constant for the five seat positions.

The method works very well for the virtual microphone located close to the driving source, there are no spikes or jumps in the resulting TFs, just a small offset at some frequencies of  $\approx 1$  dB. Due to the fact that the environment contributions to these TFs are very small -advantage of using near field sources- the measured TFs at CP do not show any sharp peaks or depths which allows to calculate a relatively smooth  $V_{k_v|k_p,l}(\omega)$ .

However, as soon as one moves to the other ears, the TFs become more complex and inaccuracies start showing up, for example at the passenger’s left ear, Pos = +12 cm at 445 Hz or for the Driver’s right ear, at Pos = -12 cm at 675 Hz among many other examples. Evidence of the fact that these anomalies are due to the calculation of  $V_{k_v|k_p,l}(\omega)$  in Eq. (4.2) can be seen if one compares the TFs at Pos = 0 with the location of the spikes at the other positions, there is a correlation between attenuations at Pos = 0 and spikes at  $\pm 6$  cm and  $\pm 12$  cm (*e.g.* in the passenger’s left ear at 445 Hz).

In order to overcome these anomalies, the aforementioned smoothing of the  $V_{k_v|k_p,l}(\omega)$  was proposed. On the one hand, in Fig.4.14, most of the anomalies present at positions  $\pm 6$  cm and  $\pm 12$  cm are reduced or in most-cases even removed. On the other hand, at Pos = 0, where the values of  $V_{k_v|k_p,l}(\omega)$  were taken and the TF should be perfect it is possible to observe that the smoothing procedure is modifying the TFs.

To optimise the visibility, *i.e.* simplify the analysis, only the amplitude responses of the TFs are shown in Fig.4.14, but the same behaviour was observed when analysing the phase-responses. Moreover, it is possible from to conclude Fig. 4.14 that the virtual-microphones method works “well” or that it gives a good “approximation” but there are some differences between the physical TFs and the virtual ones -even with the smooth procedure propose. Nevertheless, it is necessary to evaluate if the PSZ is sensitive to these differences or if the virtual method allows to achieve the performance of the system as if more microphones were present in the cabin.

Therefore, the studied metrics are plotted for all the configurations listed in Table 4.2. Additionally, the same regularisation criteria as the previous section is used, limiting the condition number to 3000.

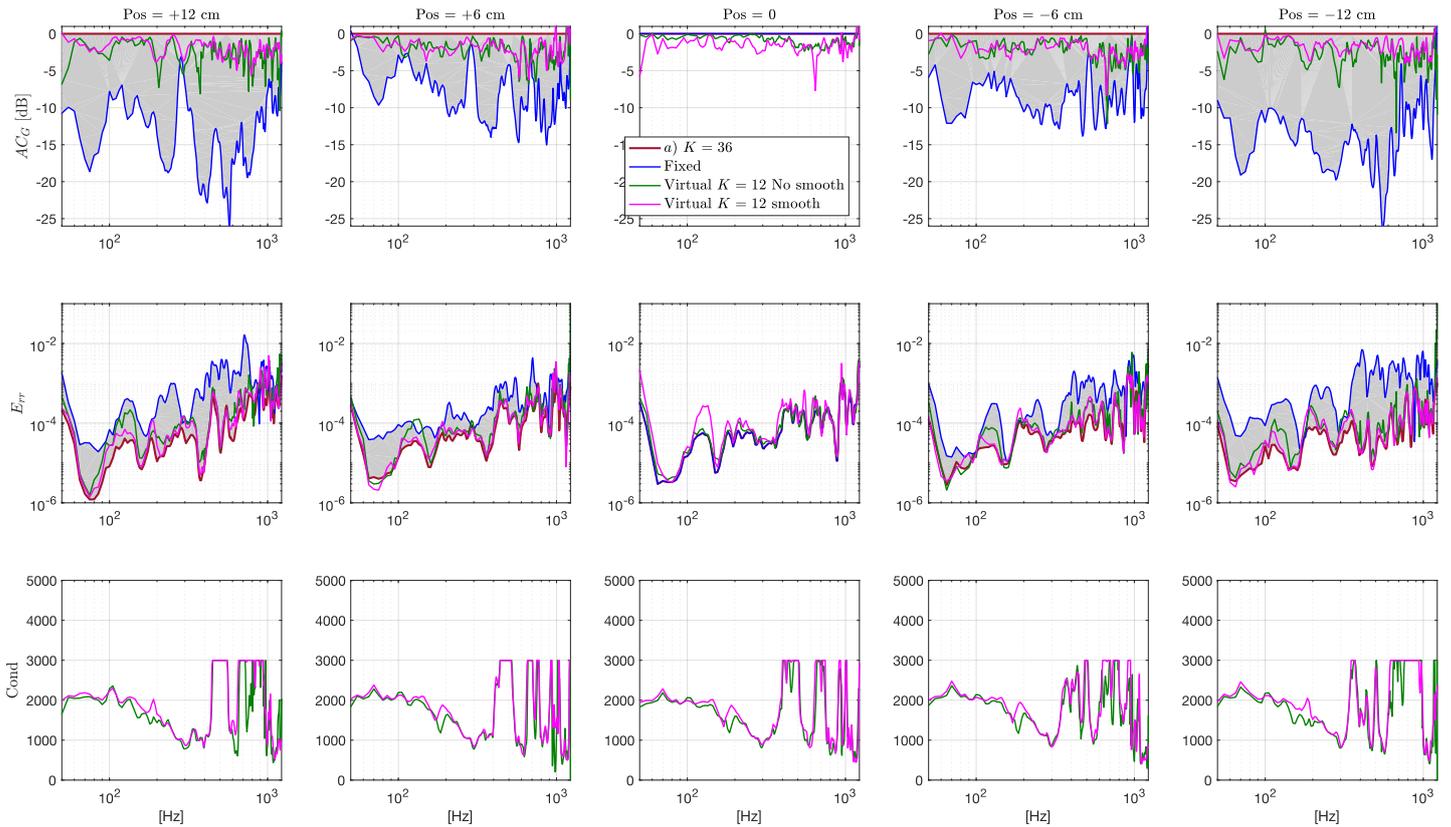


Figure 4.15: Resulting  $AC_G$ ,  $E_{rr}$  and condition number for symmetric seat positions. Results employing the re-measurement solution, using virtual microphones with and without smoothing.

From Fig. 4.15 one can draw several conclusions, firstly, it is possible to extend the number of microphones from  $K = 8$  to  $K = 12$  using the principles from the virtual sensing algorithms.

Due to the fact that to obtain the values of  $V_{k_v|k_p,l}(\omega)$  it is necessary to make the ratio between TFs, it was showed in Fig. 4.14 that one is susceptible to inaccuracies in the calculation of the TF. On the one hand, it is possible to observe that the proposed smoothing improves the performance of the system at  $\pm 12$  cm and it holds it relatively constant at  $\pm 6$  cm.

However, when analysing Pos = 0 the resulting performance is deteriorated when smoothed TFs are used. As the CP is taken at this position, the values of  $V_{k_v|k_p,l}(\omega)$  are a perfect estimation, *i.e.* at this position the smoothing means a loss of accuracy (as it was showed when analysing Fig. 4.14). Nevertheless, this only happens when the conditions of the CP are ideally re-met, which in practice is highly unlikely to happen.

### 4.5.3 Re-measurement conclusions

Firstly, one can conclude that it is possible to considerably reduce the number of microphones, from  $K = 36$  to  $K = 8$  and still obtain comparable performances. In all five positions it was possible to outperform the “Fixed” approach and generate similar  $AC$  as the  $K = 36$  with almost identical  $E_{rr}$ .

It was showed that in order to reduce the number of microphones it is necessary to take into consideration the numerical limitations of the calculation (maintain an overdetermined system and avoid ill-conditioning) and the information that the microphones are providing (which location).

The microphone the closest to the loudspeaker (microphone 7 in Fig. 4.11) is present in all the outlines with a low- $E_{rr}$  and the microphone the furthest from it (microphone 9 in Fig. 4.11) is always present in all the outlines with better  $AC_G$ .

In order to put the results of this section in perspective, one proceeded to compare its “highlights” listed in Table 4.3.

Id	$K$ Total	$K_{Phys}$	$K_{Virt}$	$k_p$	$k_v$
a)	12	12	0	4, 7 and 9	—
b)	8	8	0	7 and 9	—
c)	12	8	4	7 and 9	4
Fixed	0	0	0	—	—

Table 4.3: Microphones for comparison.

The aim of this analysis is to put in perspective the results of the microphone reduction ( $K = 12$  and  $K = 8$ ) and the results obtained from the addition of four virtual microphones against the results of  $K = 36$ .

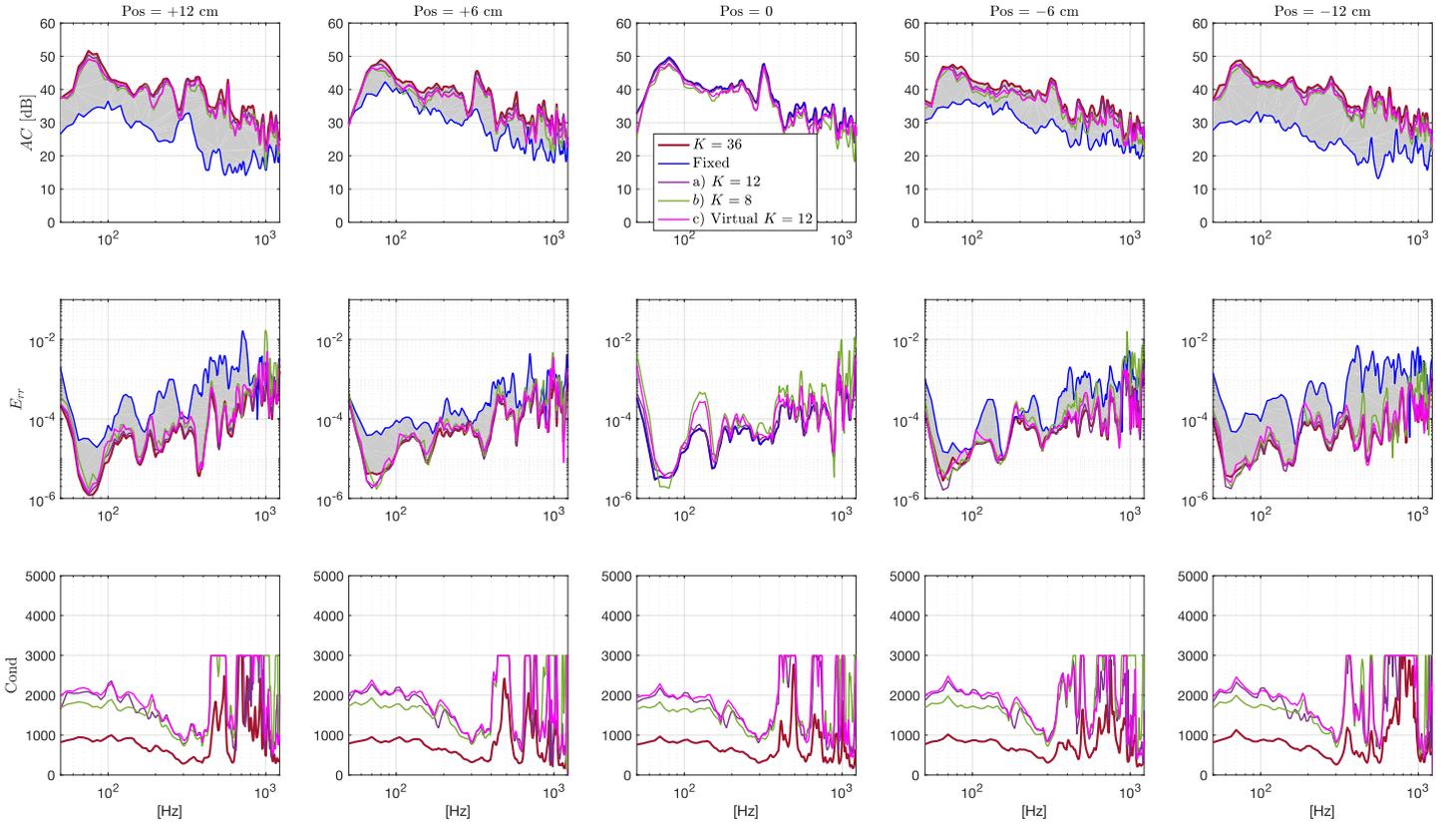


Figure 4.16: Resulting  $AC$ ,  $E_{rr}$  and condition number for symmetric seat positions. Re-measurement solutions summary.

Firstly, it is possible to observe that all three proposed methods are able to achieving comparable performance to the ideal  $K = 36$ . Although all the proposed highlights perform very well, their frequency-dependent behaviour, in addition to five seats positions makes it necessary to make some sort-of global analysis.

In addition to the  $AC_G$  (in Eq. 4.4), one could analogously define the  $E_{rrG}$  Gain as in

$$E_{rrG}|_i(\omega) = E_{rr}|_i(\omega) - E_{rr}|_{K=36}(\omega) \quad (4.4)$$

being  $i) = a), \dots, c)$ , *i.e.* the different studied configurations (Table 4.3). From Eq. 4.1 a negative  $AC_G$  means that the configuration  $i)$  achieves smaller  $AC$  than the reference one. However, a positive  $E_{rrG}$  indicates a deterioration in the performance. Note that both deterioration indicators are frequency

dependent, therefore to calculate a global indicator their mean value between 50 Hz and the Nyquist frequency are used (1225 Hz).

Mean losses [50Hz – 1225Hz]											
Id	K	Pos = +12 cm		Pos = +6 cm		Pos = 0 cm		Pos = -6 cm		Pos = -12 cm	
		$AC_G$	$E_{rrG} (\times 10^{-5})$	$AC_G$	$E_{rrG} (\times 10^{-5})$	$AC_G$	$E_{rrG} (\times 10^{-5})$	$AC_G$	$E_{rrG} (\times 10^{-5})$	$AC_G$	$E_{rrG} (\times 10^{-5})$
a)	$12_{Phys}$	-1.5	9	-1.05	7.2	-1.04	3.9	-1.31	-4.2	-1.18	0.9
b)	$8_{Phys}$	-2.98	80.6	-2.98	21.4	-3.12	76.8	-3.05	76.5	-2.78	17.3
c)	$8_{Phys} + 4_{Virt}$	-2.2	21.6	-2.02	12.1	-1.57	5.8	-2.23	-2.3	-1.78	2.6
Fixed	0	-13.54	162	-8.57	50.2	0	0	-8.23	61.3	-13.44	161

Table 4.4: Global results compared to  $K = 36$  for five symmetric positions.

Analysing the global results in Table 4.4, firstly it is possible to observe that the global deterioration of the fixed approach -using the filters from the CP everywhere- result in a much stronger deterioration than any of the other methods, at Pos  $\pm 6$  cm  $AC_G$  of  $\simeq -8$  dB and  $E_{rrG}$  of  $\simeq 5 \times 10^{-4}$ , additionally when the seats are placed at Pos  $\pm 12$  cm this configuration achieves even larger deterioration  $AC_G \simeq -13$  and  $E_{rrG} \simeq 1.6 \times 10^{-3}$ . Therefore the same conclusion (also observed in Fig. 4.16) is that all three proposed methods are capable of considerably improving the fixed approach.

On the other hand, as observed in Sec.4.5.1, configuration *a*) achieves the closest values to the  $K = 36$  configuration, followed by the configuration *c*) using the virtual microphones and in last place configuration *b*) when using  $K = 8$ .

Moreover, although when comparing the global  $AC_G$  with and without the virtual microphones there is relatively small improvement, the main gain is observed in the  $E_{rrG}$  in all the positions. Nevertheless, in the author's opinion it is difficult to make a statement regarding if these levels are detectable without a psychoacoustic evaluation.

It is important to highlight that at Pos = -6 cm, configurations *a*) and *c*) achieve negative  $E_{rrG}$ , which means that they result in even smaller  $E_{rr}$  than the  $K = 36$  configuration. This conclusion should be taken lightly, one should not forget that the PSZ problem is twofold, the  $E_{rr}$  might be lower, but there is a deterioration of  $AC$ .

Following, a summary of the advantages and disadvantages of the re-measurement solution:

#### Advantages:

- **Performance:** System is capable of maintaining both the  $AC$  as well as the  $E_{rr}$  with a small number of microphones (ideally  $K = 12$  but also  $K = 8$ ).
- **Less storage:** Only if virtual microphones are used it is necessary to only save the TF between microphones (all the necessary combinations of Eq. 4.2), otherwise no extra storage is needed.
- **Short-calibration:** It is only necessary to calibrate if virtual microphones are used, but it has been shown that with one CP suffices.
- **No anticipation:** There is no need to know in advance any of the possible EP.

#### Disadvantages:

- **Installed microphones:** It is necessary to install microphones in the cabin, with all their associated material (converter, cables, signal processing, etc.).
- **Re-measure:** Under a loss in performance, it would be necessary to stop the system, re-acquire all the loudspeaker/microphone pairs. Which might not be appealing from an industrial point of view, considering the unpleasantness of the test signals (white noise, sweep, etc.) and the time that takes to measure all  $L$  sources.
- **Re-calculate:** Once all the IRs are acquired, it is necessary to provide the computing power to recalculate the filters (with regularisation if needed) which can be time consuming and subjected to inaccuracies in the acquisition of the IRs, *e.g.* because of noise [128].

There is huge room for improvement for the re-measurement solutions. For example, a thorough study considering the influence of noise in the measured IRs must be performed. It might be necessary to incorporate an effort regularisation term based on the estimated noise in the TF to provide more stable results in the presence of noisy measurements -as proposed in Ref. [128].

In addition, due to the fact that microphone 7 is so close to the loudspeaker (in Fig. 4.11), it could be possible to assume that its TF will be constant regardless the seat positioning (allowing to have a  $K = 4$  system). In addition, the acquisition problems might be possible to overcome or improve implementing simultaneous recording of IRs, as proposed in Ref. [162] for example.

Regarding virtualisation, more complex selection processes could also be used, for example, instead of simply choosing the microphones by pairs, the contribution of several microphones could be used to obtain the virtual. In addition, it could be possible to generate a database of which microphone/CP is better for which virtual microphone, perform a more detailed smoothing of  $V_{k_v|k_p,l}(\omega)$ , etc. Additionally, the use of virtualisation algorithms is a widely-studied topic in the ANC and there might be other algorithms more suited for this application [157].

All the aforementioned improvements are out of the scope of this study, but they can lead to significant improvements. Nevertheless, before attempting to design more complex algorithms or a more complex calibration procedure, it would be necessary to evaluate the optimal microphone positioning in a more realistic prototype in order to evaluate the real limitations of these methods. For example, it has been shown the important contribution of microphone 9 (see Fig. 4.11), however, it is necessary to evaluate if it is possible to obtain similar performances with a microphone incorporated to the headrest.

## 4.6 Forced-pressure matching

In order to take one step further the use of the two previous solutions in a real application, a “re-thinking” of the classical PM/BACCPM is introduced, the forced pressure matching method (FPM). In the case of the database solution it reduces the amount of storage needed (Sec. 4.4) and in the case of the re-measurement solution (Sec. 4.5) it reduces the computational cost of the re-calculation of the filters.

Below, one goes through the development of the method, its idea behind it and an example of implementation in the car-application problem. The following development is based on the use of a deterministic approach in the frequency domain because one is assuming that the calculation time taken by the algorithm to obtain the filters is a key factor especially for the re-measurement solution (Sec. 4.5). However, there are at least three other ways for which the same principles introduced below should be applicable: in the time domain [97], using rational matrices [21] or even using an iterative approach (not a deterministic method) as in Ref. [109].

The following development is based on the 2020’s publication in the Journal of the Audio Engineering Society [163].

### 4.6.1 Forced Pressure Matching theory

In order to explain the theory behind the FPM, one is going to go through the calculation of the filter combinations using the deterministic approach again as in Sec. 2.3.1, but with a small twist. From the aforementioned section, one knows that it is possible to express the pressures for the  $K$  microphones, for an angular frequency  $\omega$  as

$$\hat{\mathbf{p}}(\omega) = \hat{\mathbf{G}}(\omega)\hat{\mathbf{q}}(\omega), \quad (4.5)$$

being

$$\hat{\mathbf{q}}(\omega) = [\hat{q}_1(\omega), \hat{q}_2(\omega), \dots, \hat{q}_L(\omega)]^T \quad (4.6)$$

and  $\hat{\mathbf{G}}$  the  $(K \times L)$  matrix of complex transfer functions (just a copy of Eqs. 2.11 and 2.12). The  $(\omega)$  term is ignored from this point forward for compactness.

In the classical PSZ approaches, every filter  $\hat{q}$  is considered as a unknown variable that needs to be calculated. However, there are some cases where it might not be necessary to calculate every filter. For

example, if a specific response wants to be applied to one source (or more than one) or if the response of some filters is known in advance.

Hence, assuming that among  $\hat{\mathbf{q}}$  there are  $N$  weights ( $\hat{\mathbf{q}}_f$ ) one wants to force and  $L - N$  weights  $\hat{\mathbf{q}}_c$  to calculate, the vector of complex weights from Eq. 4.6 can be expressed as:

$$\hat{\mathbf{q}} = \begin{bmatrix} \hat{\mathbf{q}}_c \\ \hat{\mathbf{q}}_f \end{bmatrix} = \{[\hat{q}_{c1}, \hat{q}_{c2}, \dots, \hat{q}_{c(L-N)}][\hat{q}_{f1}, \hat{q}_{f2}, \dots, \hat{q}_{fN}]\}^T, \quad (4.7)$$

where  $0 < N < L$ ; otherwise if  $N = 0$  is the classical deterministic PM approach and if  $N = L$  all the filters are known.

Furthermore, as done for  $\hat{\mathbf{q}}$ , in  $\hat{\mathbf{G}}$  there are some TF that correspond to the sources associated to the forced filters (denoted with the subindex 'f') and other that correspond to the filters to calculate (denoted with the subindex 'c'), as in

$$\hat{\mathbf{G}} = \begin{bmatrix} \hat{\mathbf{G}}_c & \hat{\mathbf{G}}_f \end{bmatrix}, \quad (4.8)$$

where  $\hat{\mathbf{G}}_c$  is the  $K \times (L - N)$  matrices of complex TF that correspond to the complex weights to calculate and  $\hat{\mathbf{G}}_f$  is the  $K \times N$  matrices of complex TF that correspond to the forced complex weights.

Introducing Eqs. 4.7 and 4.8 in Eq. 4.5 the pressure in the  $K$  microphones can be expressed as

$$\hat{\mathbf{p}} = \begin{bmatrix} \hat{\mathbf{G}}_c & \hat{\mathbf{G}}_f \end{bmatrix} \begin{bmatrix} \hat{\mathbf{q}}_c \\ \hat{\mathbf{q}}_f \end{bmatrix}. \quad (4.9)$$

Moreover, we defined the error in the  $K$  microphones as the difference between the pressure and a desired pressure, as in Eq. 2.25 the error can be expressed as

$$\hat{\mathbf{e}} = \hat{\mathbf{p}}_d - \begin{bmatrix} \hat{\mathbf{G}}_c & \hat{\mathbf{G}}_f \end{bmatrix} \begin{bmatrix} \hat{\mathbf{q}}_c \\ \hat{\mathbf{q}}_f \end{bmatrix}, \quad (4.10)$$

being  $\hat{\mathbf{p}}_d$  is the  $K \times 1$  vector of desired pressures (also known as the target pressures).

Furthermore, the resulting pressure due to the forced complex weights can be calculated as follows

$$\hat{\mathbf{p}}_f = \hat{\mathbf{G}}_f \hat{\mathbf{q}}_f \quad (4.11)$$

and subtracted from the original desired pressures as

$$\hat{\mathbf{e}} = [\hat{\mathbf{p}}_d - \hat{\mathbf{p}}_f] - \hat{\mathbf{G}}_c \hat{\mathbf{q}}_c. \quad (4.12)$$

Hence, in Eq. 4.12 a new vector of desired pressures can be defined ( $\hat{\mathbf{p}}_{dc}$ ) containing both the information from the original  $\hat{\mathbf{p}}_d$  and the information from the forced filters. Therefore, the error becomes

$$\hat{\mathbf{e}}_c = \hat{\mathbf{p}}_{dc} - \hat{\mathbf{G}}_c \hat{\mathbf{q}}_c. \quad (4.13)$$

As in Sec. 2.3.3, the cost function of the Forced pressure matching -in the frequency domain- can be expressed as

$$\hat{\mathcal{L}}_c = \hat{\mathbf{e}}_c^H \hat{\mathbf{e}}_c = (\hat{\mathbf{p}}_{dc} - \hat{\mathbf{G}}_c \hat{\mathbf{q}}_c)^H (\hat{\mathbf{p}}_{dc} - \hat{\mathbf{G}}_c \hat{\mathbf{q}}_c) \quad (4.14)$$

To be more coherent with the PSZ literature it is possible to separate the cost function (Eq. 4.15) according to the microphones that correspond to the  $BZ$  and  $DZ$  as

$$\hat{\mathcal{L}}_c = (\hat{\mathbf{p}}_{dcBZ} - \hat{\mathbf{G}}_{cBZ} \hat{\mathbf{q}}_c)^H (\hat{\mathbf{p}}_{dcBZ} - \hat{\mathbf{G}}_{cBZ} \hat{\mathbf{q}}_c) + (\hat{\mathbf{p}}_{dcDZ} - \hat{\mathbf{G}}_{cDZ} \hat{\mathbf{q}}_c)^H (\hat{\mathbf{p}}_{dcDZ} - \hat{\mathbf{G}}_{cDZ} \hat{\mathbf{q}}_c) \quad (4.15)$$

Lastly, following the Lagrange multipliers method as in Sec. 2.3.3, it is possible to obtain the expression of the complex weights by taking the gradient of the cost function and equating it to zero ( $\nabla \hat{\mathcal{L}}_c = \mathbf{0}$ )

$$\hat{\mathbf{q}}_c = (\hat{\mathbf{G}}_{cBZ}^H \hat{\mathbf{G}}_{cBZ} + \hat{\mathbf{G}}_{cDZ}^H \hat{\mathbf{G}}_{cDZ})^{-1} (\hat{\mathbf{G}}_{cBZ}^H \hat{\mathbf{p}}_{dcBZ} + \hat{\mathbf{G}}_{cDZ}^H \hat{\mathbf{p}}_{dcDZ}). \quad (4.16)$$

It is important to highlight, despite its redundancy, that there is no need to re-calculate the already forced weights ( $\mathbf{q}_f$ ), so the total number of complex weights (filters) to calculate as well as the size of the complex TFs matrices involved in the calculation is reduced.

### 4.6.2 The forced filters, the desired and their relationship

In Eq. 4.15 one arrived to an expression of the pressure matching cost-function when some of the filters involved in the calculation are forced and now the challenge is to choose which filters to force and how they modify the desired pressure.

As it was explained in Sec. 3.1, to achieve a transparent system, the desired pressure in the  $BZ$  is the resulting pressure when  $P_R$  and  $P_L$  are driven in phase (see Fig. 3.9) and the desired pressure in the  $DZ$  is equal to  $\emptyset$  to achieve a potential energy minimisation. Therefore, it is possible to define the complex weight vector ( $\hat{\mathbf{q}}_d$ ) used to calculate the desired as the  $L \times 1$  vector full of zeros except for the complex weights that correspond to  $P_R$  and  $P_L$  which are equal to 1 as in

$$\begin{aligned}\hat{\mathbf{q}}_d &= [\hat{q}_{dP_O}, \hat{q}_{dP_R}, \hat{q}_{dP_L}, \hat{q}_{dP_I}, \hat{q}_{dD_I}, \hat{q}_{dD_R}, \hat{q}_{dD_L}, \hat{q}_{dD_O}] \\ &= [0, 1, 1, 0, 0, 0, 0, 0]\end{aligned}\quad (4.17)$$

being  $q_{dl}$  the complex weight applied to the  $l$  source used to calculate the desired. Consequently, the desired pressure can be expressed as

$$\mathbf{p}_d = \begin{cases} \hat{\mathbf{p}}_{dBZ} = \hat{\mathbf{G}}_{BZ} \hat{\mathbf{q}}_d \\ \hat{\mathbf{p}}_{dDZ} = \emptyset \end{cases} \quad (4.18)$$

To define the forced filters, one is going to come back to Chapter 3. As it was observed in all the reviewed levels of acoustic complexity -from an anechoic chamber to a car cabin- the filters associated to the sources  $P_R$  and  $P_L$  (the closest to the  $BZ$ ) always resulted in almost perfect  $\delta[n]$ ; also referred to as the reproduction filters. Thus, there is no need to re-calculate them every time (Re-measurement approach, Sec. 4.5) or to store them for different acoustic conditions (Sec. 4.4), so these filters are chosen to be forced in the calculation.

Consequently, the number of forced filters is  $N = 2$ , and they are forced as perfect  $\delta[n]$  functions, as one is using a frequency domain formulation

$$\hat{\mathbf{q}}_f = \begin{bmatrix} \hat{q}_{f1} \\ \hat{q}_{f2} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \quad (4.19)$$

Hence, the resulting pressure due to the forced complex weights (using Eq. 4.11) is

$$\mathbf{p}_f = \begin{cases} \hat{\mathbf{p}}_{fBZ} = \hat{\mathbf{G}}_{fBZ} \hat{\mathbf{q}}_f \\ \hat{\mathbf{p}}_{fDZ} = \hat{\mathbf{G}}_{fDZ} \hat{\mathbf{q}}_f \end{cases} \quad (4.20)$$

Then, it is possible to calculate the “new” desired pressure, from Eq. 4.18 and Eq. 4.20 as

$$\hat{\mathbf{p}}_{dc} = [\hat{\mathbf{p}}_d - \hat{\mathbf{p}}_f] = \begin{cases} \hat{\mathbf{G}}_{BZ} \hat{\mathbf{q}}_d - \hat{\mathbf{G}}_{fBZ} \hat{\mathbf{q}}_f \\ \emptyset - \hat{\mathbf{G}}_{fDZ} \hat{\mathbf{q}}_f \end{cases} \quad (4.21)$$

It is important to notice that in the  $BZ$  the forced-weights ( $\hat{\mathbf{q}}_f$ ) correspond to the exactly the same weights that the ones are being used for the definition of  $\hat{\mathbf{p}}_d$  (see Eq. 4.17), *i.e.*  $\hat{\mathbf{G}}_{BZ} \hat{\mathbf{q}}_d = \hat{\mathbf{G}}_{fBZ} \hat{\mathbf{q}}_f$ . Hence, Eq. 4.21 becomes

$$\hat{\mathbf{p}}_{dc} = \begin{cases} \emptyset \\ -\hat{\mathbf{G}}_{fDZ} \hat{\mathbf{q}}_f \end{cases} \quad (4.22)$$

Therefore, if one introduces Eq. 4.22 in the cost function (Eq. 4.15) it yields:

$$\begin{aligned}\hat{\mathcal{L}}_c &= (\hat{\mathbf{G}}_{cBZ} \hat{\mathbf{q}}_c)^H (\hat{\mathbf{G}}_{cBZ} \hat{\mathbf{q}}_c) + (\hat{\mathbf{p}}_{dcDZ} - \hat{\mathbf{G}}_{cDZ} \hat{\mathbf{q}}_c)^H (\hat{\mathbf{p}}_{dcDZ} - \hat{\mathbf{G}}_{cDZ} \hat{\mathbf{q}}_c) \\ &= \hat{\mathbf{q}}_c^H \hat{\mathbf{G}}_{cBZ}^H \hat{\mathbf{G}}_{cBZ} \hat{\mathbf{q}}_c + (\hat{\mathbf{p}}_{dcDZ} - \hat{\mathbf{G}}_{cDZ} \hat{\mathbf{q}}_c)^H (\hat{\mathbf{p}}_{dcDZ} - \hat{\mathbf{G}}_{cDZ} \hat{\mathbf{q}}_c)\end{aligned}\quad (4.23)$$

which presents exactly the opposite behaviour that in the well-known combined solution (in 2.31, or *e.g.* in Refs. [97, 125, 126, 133]); a potential energy minimisation term in the  $BZ$  and an error minimisation term in the  $DZ$ .

Then, the expression of the non-forced filters (as in Eq. 4.16) is given by

$$\hat{q}_c = (\hat{\mathbf{G}}_{cBZ}^H \hat{\mathbf{G}}_{cBZ} + \hat{\mathbf{G}}_{cDZ}^H \hat{\mathbf{G}}_{cDZ})^{-1} (\hat{\mathbf{G}}_{cDZ}^H \hat{\mathbf{p}}_{dcDZ}) \quad (4.24)$$

where it is only the desired pressure in the  $DZ$  that needs to be specified.

### 4.6.3 The FPM performance: Complexity №5 metrics

In Ref. [163], the FPM was reviewed using anechoic conditions (Complexity №2 in Sec. 3.2.2), therefore the following results can be seen as an extension of this work in a car cabin.

#### Invariant acoustics

Therefore, one proceeded to compare the performances of the PM and the FPM algorithm in a car cabin with the seats in the centre (same analysis as the one in Sec. 3.4).

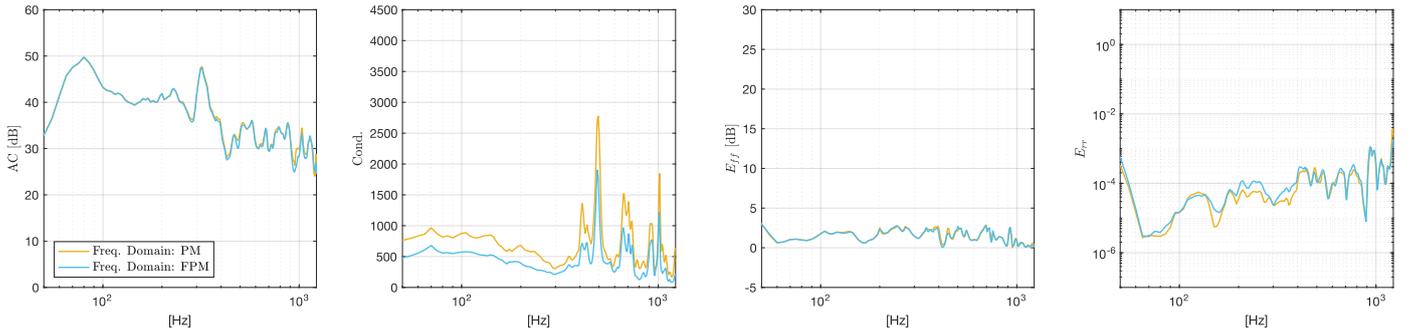


Figure 4.17: Measurements results in a car cabin at Pos = 0, pressure matching (yellow) and forced pressure matching (cyan).

As in the anechoic results published in Ref. [163], the  $AC$  and the  $E_{ff}$  obtained from both formulations present almost identical performances regardless of the algorithm used when applied in a car cabin (see Fig. 4.17). On the other hand, the FPM improves the conditioning of the problem (in some frequencies by a factor of two or more). Nevertheless both formulations are far from ill-conditioned, *i.e.*, it does not seem to be necessary to add some sort of regularisation term, maybe just a localised regularisation in the largest Cond. peaks at 495 Hz and 1015 Hz.

Moreover there is a small change in performance when it comes to the Err; the PM algorithm results in a slightly smaller reproduction error.

Lastly, the PM took 7.2 ms whilst the FPM took 5.7 ms, 20.8% faster. The calculation time is expressed as the mean value after  $2 \times 10^4$  iterations. The code was implemented under MATLAB, under Windows 7 Enterprise, using an Intel Core(TM) i7-6820HQ CPU with a RAM memory of 16 GB and a 64-bit architecture.

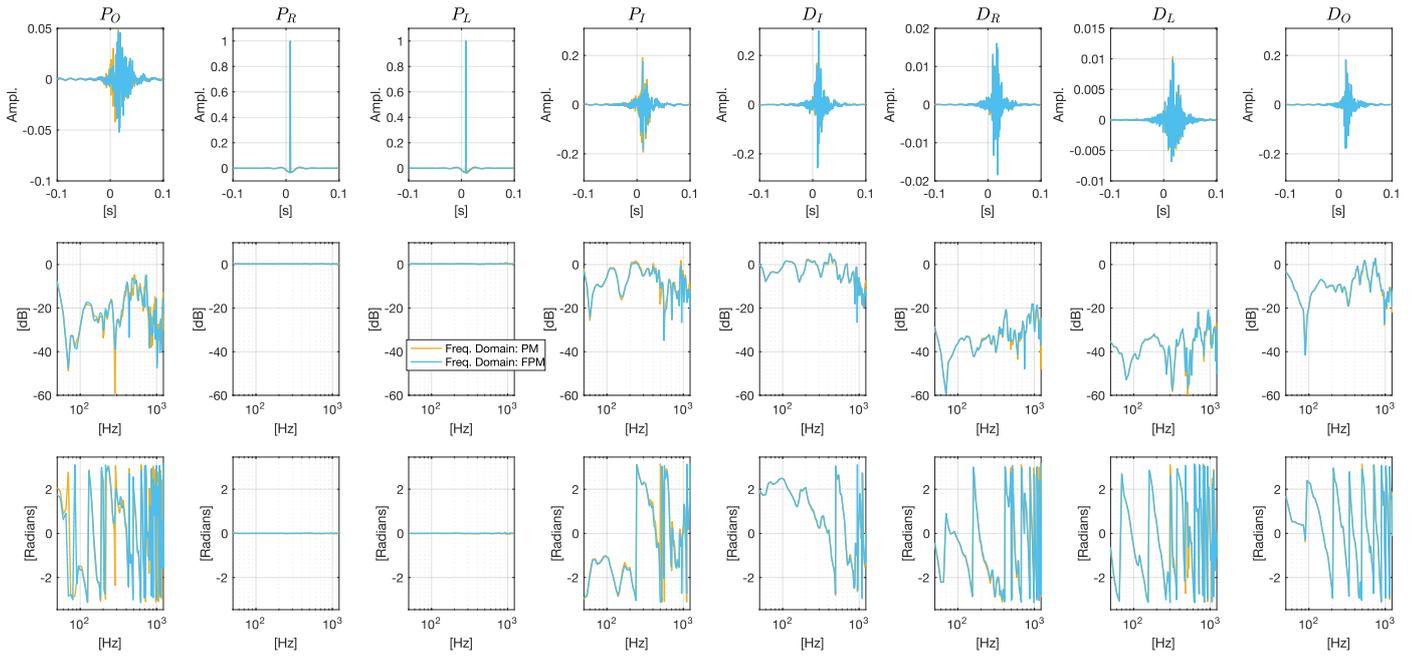


Figure 4.18: Resulting filters using the PM and the FPM formulations, employing eight loudspeakers, in a car cabin.

To evaluate the impulse responses from Fig. 4.18 it is useful to separate between reproduction and cancellation filters (as in Chapter 3).

Firstly, it is possible to observe that the reproduction filters in both formulations are very close. The variations between the methods are  $\leq 0.7$  dB and  $\leq 0.03$  radians (see Fig. 4.18). In addition, these small discrepancies can be seen only because the scale of  $P_R$  and  $P_L$  has been considerably enlarged. On the other hand, when analysing the cancellation filters one can conclude that both formulations result in similar performances. The main differences in most of these filters are in the depth of the attenuation when a very steep notch-type response is obtained (*e.g.*  $P_O$  at 285 Hz or  $P_I$  at 495 Hz). Additionally, these very steep corrections also result in a sudden phase-shift which is not always equal between the methods.

In spite of the fact that both methods are originated from the same expression of the error (Eq. 4.10) and the minimisation of this error is also done in a least-squares sense, as one is reducing the number of complex weight to calculate the solution, it results in two different methods which result in two different filter combinations.

### Sensitivity to noise

In the case of the re-measurement approach (Sec. 4.5), as it is not possible to control the conditions in which the IR are measured one key aspect that is important to analyse is the system's sensitivity to noise.

For this, a noisy-environment was emulated adding white Gaussian noise to the TF. Moreover, in spite of adding a different noise to each loudspeaker-microphone pair, the same matrix of noises is added to both methods. Four different levels of noise were used, resulting in four cases of signal-to-noise ratio ( $SNR$ ) as in table 4.5.

Case No.	Mean SNR
0	No Noise
1	35 dB
2	20 dB
3	5 dB
4	-10 dB

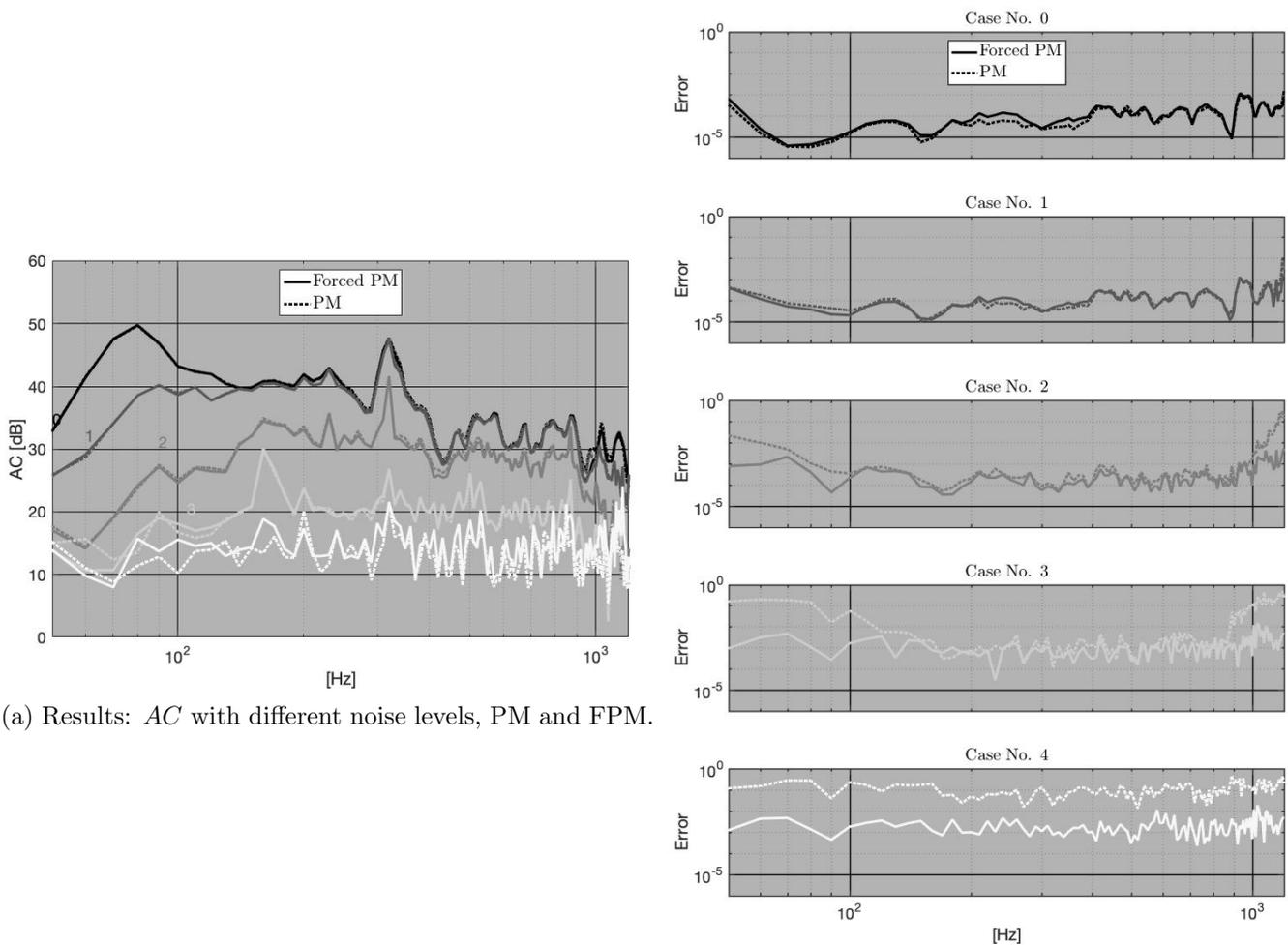
Table 4.5: Table of different  $SNR$  cases studied.

The  $SNR$  for one microphone loudspeaker combination ( $kl$ ) is defined as

$$SNR_{G_{k,l}} = \frac{\sum_0^\Omega |G_{kl}(\omega)|^2}{\sum_0^\Omega |Noise_{kl}(\omega)|^2}, \tag{4.25}$$

where  $Noise_{kl}$  is the noise introduced to the  $G_{kl}$  TF and  $\Omega$  is the length of the TFs.

It is necessary to highlight that the mean  $SNR$  is taken only as an indicator; the differences between the amplitudes of the TF is very large —some microphones are much closer to the loudspeaker than others. For example, in Case No. 1 the maximum  $SNR > 53$  dB while the minimum  $SNR < 15$  dB.



(a) Results:  $AC$  with different noise levels, PM and FPM.

(b) Results:  $Err$  with different noise levels, PM and FPM.

Figure 4.19: Comparison between classical PM and FPM introducing five different levels of noise (see Table 4.5).

Regarding the  $AC$ , the analysis of the sensitivity to noise obtained in the car cabin (Fig. 4.19a) is perfectly analogous to the one in anechoic conditions [163]. Both algorithms are equally deteriorated in the presence of noise. The deterioration of the  $AC$  is directly proportional to the  $SNR$ , evidence of this can be seen comparing Case No.0 to Case No.1 and the small deterioration in the lowest part of the spectrum where worse  $SNR$  are present.

When analysing the  $E_{rr}$ , although this results were oriented towards comparing two methods, the designed system on its own presents a very good sensitivity to noise. More specifically, it is necessary to use relatively small  $SNR$  levels in order to deteriorate its performance regardless the formulation, this is a direct consequence of the use of near-field sources.

As it was observed in Ref. [163] and in Fig. 4.19b, when no -extra- noise is present, the PM slightly outperforms the FPM method. In the  $SNR$  Case No.1, it is possible to observe that both formulations are performing almost identically (the FPM slightly better in the lowest part of the spectrum). Nevertheless, in Case No.2 one can already distinguish that the PM's performance is being deteriorated at the extremes of the studied bandwidth (especially below 100 Hz) and that the FPM results for the complete bandwidth in a smaller or equal  $E_{rr}$ . And in already smaller levels of  $SNR$ , (Case No.3 and above), the FPM is capable of maintaining a "minimum" level of reproduction whilst the PM keeps on deteriorating.

#### 4.6.4 The FPM performance: Extension to a variable car cabin

Lastly, one proceeded to calculate the PSZ filters when using the FPM and to compare it to the original PM using the 5 symmetric seat positioning (as in Fig.4.1a).

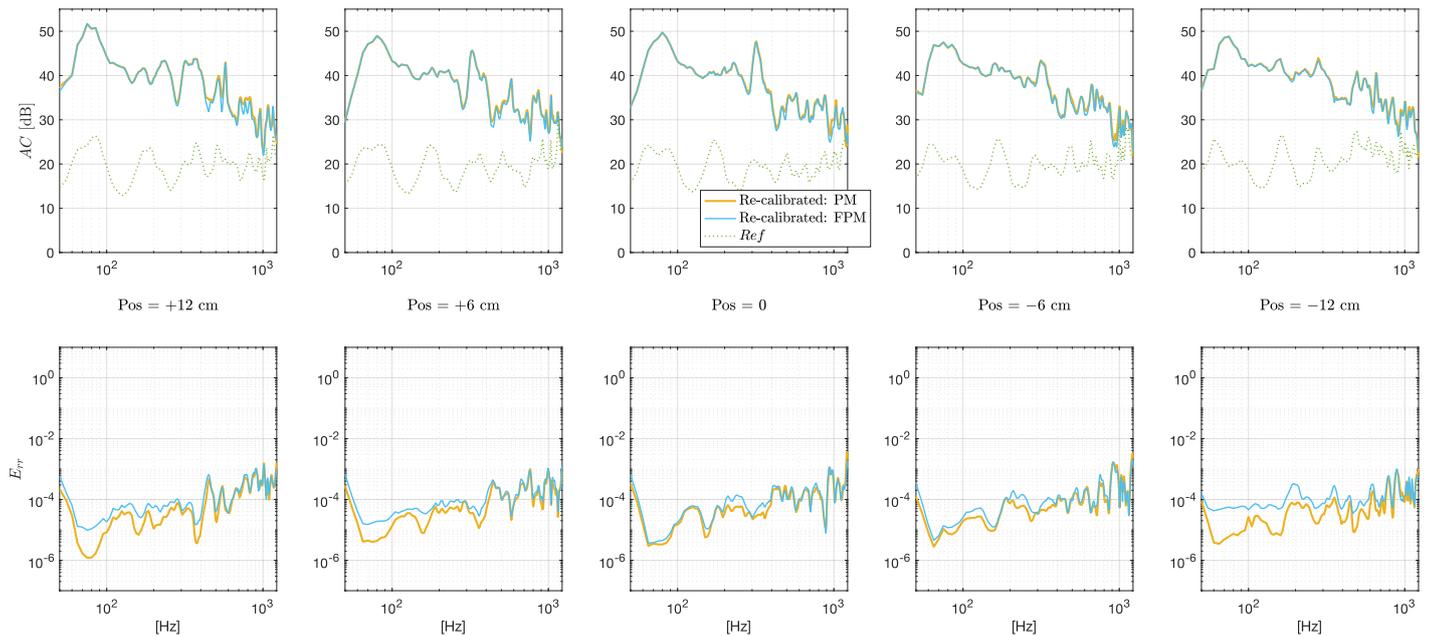


Figure 4.20: Resulting  $AC$  and  $E_{rr}$  for symmetric seat positions, using the PM and FPM.

In all five seat-positions, both the PM and FPM result in almost identical  $AC$  regardless the algorithm used (Fig. 4.20). Regarding the  $E_{rr}$ , it is possible to observe that the PM algorithm results in slightly better performance (the largest deterioration is of the order of  $4.3 \times 10^{-5}$ ).

### 4.6.5 Discussion FPM

The FPM was introduced as an alternative of the PM methods when specific filter responses are forced into the calculation. This section showed its analytical demonstration and extended the results obtained in Ref. [163] to a car cabin.

When comparing the performance of the PM and the FPM, both result in the same  $AC$  and very similar  $E_{rr}$ . The FPM not only allows to have a specific response in some filters, but it can also be used to reduce the number of variables involved in the calculation which leads to a faster calculation and smaller sensitivity to noise.

The FPM can be thought as the PM that has been already fed with a part of the solution, in this application the reproduction part of the problem. Moreover, one can expand from Eq. 4.10 the TF matrix in two, the TFs corresponding to the microphones in the  $BZ$  and the ones corresponding to the microphones in the  $DZ$  as

$$\hat{\mathbf{e}} = \begin{bmatrix} \hat{\mathbf{p}}_{dBZ} \\ \hat{\mathbf{p}}_{dDZ} \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{G}}_{BZc} & \hat{\mathbf{G}}_{BZf} \\ \hat{\mathbf{G}}_{DZc} & \hat{\mathbf{G}}_{DZf} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{q}}_c \\ \hat{\mathbf{q}}_f \end{bmatrix}. \quad (4.26)$$

Furthermore, one can also separate Eq. 4.26 between the contributions from the forced filters ( $\hat{\mathbf{q}}_f$ ) and the filters to calculate ( $\hat{\mathbf{q}}_c$ ) to each zone, therefore from Eq. 4.26 one can obtain

$$\hat{\mathbf{e}} = \hat{\mathbf{p}}_{dBZ} - \left[ \underbrace{\hat{\mathbf{G}}_{BZf} \hat{\mathbf{q}}_f}_{\text{(a)}} + \underbrace{\hat{\mathbf{G}}_{BZc} \hat{\mathbf{q}}_c}_{\text{(b)}} \right] + \hat{\mathbf{p}}_{dDZ} - \left[ \underbrace{\hat{\mathbf{G}}_{DZf} \hat{\mathbf{q}}_f}_{\text{(c)}} + \underbrace{\hat{\mathbf{G}}_{DZc} \hat{\mathbf{q}}_c}_{\text{(d)}} \right]. \quad (4.27)$$

In Eq. 4.27, one can observe a summation of six “different pressures” of the posed problem, two desired pressures, and four resulting pressures. One is going to analyse each of the different resulting pressures which represent a different part of the studied problem:

- (a) Near-field reproduction: this resulting pressure corresponds to the reproduction part of the problem, *i.e.* the pressure generated by  $P_R$  and  $P_L$  in the near-field of the  $BZ$ .
- (b) Deterioration of the reproduction: this resulting pressure can be seen as the deterioration of the reproduction due to all the calculated filters (referred to as cancellation filters after Chapter 3).
- (c) Crossed-zones term: this corresponds to the pressure generated by the reproduction sources/filters ( $P_R$  and  $P_L$ ) in the  $DZ$ .
- (d) Attenuation: this is solely the pressure generated in the  $DZ$  by all the cancellation sources.

If one introduces the information of the forced filters and the desired from Sec. 4.6.2, Eq. 4.27 becomes

$$\hat{\mathbf{e}} = - \underbrace{\hat{\mathbf{G}}_{BZc} \hat{\mathbf{q}}_c}_{\text{(b)}} - \left[ \underbrace{\hat{\mathbf{G}}_{DZf} \hat{\mathbf{q}}_f}_{\text{(c)}} + \underbrace{\hat{\mathbf{G}}_{DZc} \hat{\mathbf{q}}_c}_{\text{(d)}} \right]. \quad (4.28)$$

From the four resulting pressures in Eq. 4.27, only three remain after the forced filters (Eq. 4.28), due to the fact that (a) is no longer in the equation, it is possible to state that the forced filters “solved” the reproduction part of the error. However, this does not mean neither that the  $BZ$  can be ignored ((b) still in the equation) nor that the contribution of the forced filters can be ignored ((c) also in the equation). For example, if (b) was not present, the problem would be only the generation of a zone of quietness [1], however, it is in the requirements of the application that it is necessary to control both zones.

The fact that the expression of the error is modified, implies that the system’s cost-function is modified as well. In this application, the FPM’s cost function results in an “inversion” of the original combined solution cost function (Sec. 4.6.2). Moreover, the  $DZ$  results in an error minimisation term cancelling what

is coming from the forced loudspeakers (hence the “negative” sign in the  $\mathbf{p}_{dc}$  in Eq. 4.22) and it needs to do this without disturbing the  $BZ$  (the potential energy minimisation term in Eq. 4.23).

The main drawback of the FPM is that it is necessary to have some knowledge of the solution, *i.e.* the “good solution” needs to be fed to the algorithm in order to profit all the advantages of the FPM reviewed on this thesis. Otherwise, the forcing acts as an imposition to the algorithm and it can severely affect its performance.

In other words, the FPM can be also seen as a more constrained PM that results in a reduction/simplification of the “degrees of freedom” that the algorithm has to solve (the size of  $\mathbf{q}$  is by definition bigger than the size of  $\mathbf{q}_c$ ). Evidence of this can be seen in Fig. 4.18 from the differences in the calculated filters. Moreover, the effect of this simplification can be the explanation of why the FPM is better conditioned than the PM (Fig. 4.17).

The FPM was explained in this section using the Frequency domain formulation, where the complex weights are expressed for a single frequency and needs to be repeated for the bandwidth one wants to control. Nevertheless, the FPM principle -subtracting the contribution of the forced filters from the desired (Eq. 4.12) and reducing the calculation matrices accordingly- can be also applied to other formulation based on an error minimisation, *e.g.* the BACCPM.

The fact that  $\mathbf{G}_c$  is smaller than  $\mathbf{G}$  and to have less filters to calculate  $\mathbf{q}_c$  is  $N$  times smaller than  $\mathbf{q}$  has a direct reduction on the computational cost. For example in this application, if a database solution is chosen (Sec. 4.4), the FPM allows to reduce the number of stored filters by a 25% because the filters associated with the sources  $P_R$  and  $P_L$  are always  $\delta[n]$  functions. On the other hand, if a re-measurement approach is chosen (Sec 4.5), the FPM reduces the time of calculation of the filters by  $\approx 21\%$ .

Furthermore, a more thorough analysis could be performed to fully evaluate the effect of the loudspeaker reduction on the computational cost, especially because the frequency domain formulation is inherently so fast that the accuracy of the time measurement already has an embedded error. In addition, taking as a reference the work presented in [21], one could also add to the calculation time, the working memory requirements and use different array sizes, different filter lengths, different formulations, etc. But this is out of the scope of this analysis.

Lastly, it would be interesting not only to apply the FPM when a piece of the answer is known (as done in this study), but also to use the filter forcing as a design constraint. For example, as the frequency band of interest of this report is up to 1 kHz, it is assumed that the higher part of the hearing spectrum is handled by a different technology (see Chapter 2), therefore, it might be interesting to incorporate into the formulation a future crossover as a forced filter in the optimisation.

## 4.7 Summary

The main goal of this chapter was to study the PSZ system, in the presence of acoustical-modifications in the car cabin.

At the beginning of the chapter, one proceeded to extend the literature review (from Chapter 2) incorporating the publications where the sensitivity of the PSZ was studied under different modifications in the acoustic environment.

Secondly, one reviewed the acquisition of a broad library of IRs to evaluate the system’s performance when the seat-position changes, the headrest-position changes and when passengers are added in the rear seats.

Subsequently, one proceeded to evaluate the system under different acoustic conditions, both employing a fixed set of filters (from one specific calibration position) and re-calculating the filters after each modification. Although the system is capable of achieving the aimed performance under all the different conditions, a specific filter combination needs to be applied for each acoustic condition.

Therefore two solutions were proposed directly derived from the deterministic approaches: the database and the re-measurement one.

The data-base solution is based on the anticipation of every modification in the acoustic environment that might occur in a car cabin. The re-measurement solution is based on the re-characterisation of the acoustic environment in the presence of a loss of performance.

Lastly, in order to make both these approaches more attractive for the car application, one introduced the forced-pressure matching formulation. The use of the FPM for the application in hand, reduces the amount of storage needed by the database-solution by 25% without compromising the performance. Additionally, if a re-measurement solution is used, it reduces the calculation time taken by the algorithm by  $\simeq 21\%$  without compromising the performance.

### **The balance of the chapter**

The main contributions of the chapter are listed below.

- 1) Acquisition of an extensive data base of acoustic modifications in a car cabin.
- 2) Headrest PSZ system, validated under different acoustic conditions: seat positioning, headrest positioning and addition of passengers in the rear seats.
- 3) Study on the microphone reduction.
- 4) Extension of the number of microphones implementing a virtualisation-like technique.
- 5) Introduction of the forced pressure matching algorithm.
  - Patent pending: N°FR2005563
  - Published in the Journal of the audio engineering society [163]



## Chapter 5

# Adaptive approach

In chapter 2, the use of a deterministic approach to calculate the optimal filter combination was explained. The main disadvantage of these deterministic methods is that it is assumed that the IRs used to calculate the filters perfectly represent the system and that they do not change over time.

Throughout Chapter 4, it was showed that “common” modifications in the car -like the position of the seat, the headrest or the addition of passengers in the rear seats- deteriorate the generation of PSZ, and exposed the need of a dedicated filter combination for each source/environment/zones combination (see block diagram in Fig. 2.13). Furthermore, two strategies finding the way around this limitation of deterministic approaches were analysed and a novel way of posing the error minimisation algorithms (forced-pressure matching) was introduced in order to make these solutions more attractive to the automotive industry. Nevertheless, either one has the capabilities of anticipating and storing all the solutions or one needs to stop the entertainment system, run a re-acquisition of the IRs and re-calculate the filters.

In this chapter, as an alternative to the deterministic approach, the use of adaptive filter algorithms is proposed to overcome the acoustic-variability of the car cabin. The application of iterative/adaptive approaches<sup>1</sup> to generate PSZ has started to gain more attention in the last few years, but it is still is a topic that has not had a lot of development despite the fact that the clues have been there since the beginning [1].

Therefore, this chapter is constructed as follows: firstly an extension of the literature review focused on iterative algorithms for PSZ, secondly a review of the theoretical development of the FxLMS, thirdly, a parametric study of the FxLMS. With the conclusions of previous sections, one considered the possibility of reducing the number of microphones, the extension of the FPM principle to the FxLMS algorithms and lastly, an implementation of previous ideas considering different types of acoustic changes.

### 5.1 Literature review extension: PSZ and the adaptive algorithms

The use of iterative approaches to calculate the PSZ filters has been overlooked or dismissed in several occasions despite that in the PSZ’s foundational paper -Druvesteyn and Garas 1997 [1]- the method proposed to control the lowest part of the spectrum was the widely-studied recursive filtered-x least-mean-squares method (FxLMS). Although their focus was only to generate a zone of silence, *i.e.* only to control the *DZ*, it was an iterative approach applied for the generation of PSZ.

Before digging into the topic, the author would like to mention that the adaptive-filter algorithm literature is vast. Researchers have been publishing on the topic for at least 60 years, for example the first formulation of the LMS algorithm dates back to the year 1960 [164]. Hence, the aim of this chapter is not to innovate on the adaptive-filter domain but to take advantage of their findings and apply them to the PSZ problematic.

---

<sup>1</sup>In this dissertation, the terms iterative approach and adaptive approach are used as synonyms. These can also be referred to as self-designing algorithms, recursive algorithms, etc. [124]

The core idea of the iterative/recursive methods, is to improve the impulse responses of the filters ( $\mathbf{w}$ ) at an iteration  $n + 1$  by adding the update terms  $\mathbf{Y}$  which are computed at the previous iteration  $n$  [109] as in

$$\mathbf{w}[n + 1] = \mathbf{w}[n] + \mathbf{Y}[n]. \quad (5.1)$$

The algorithm starts from some predetermined set of initial conditions ( $\mathbf{w}[0]$  usually defined as zeros  $\mathbf{0}$ ) representing whatever we know about the environment and after successive adaptation cycles the algorithm converges to the optimum solution in some statistical sense [124]. It is necessary to mention that in this thesis it is assumed that the update happens every sample  $[n]$ , however, it is also possible to use block-adaptive filters (calculate  $\mathbf{Y}$  from a block of  $[n]$ ) [124, 165, 166, 167].

The iterative approaches assume:

- It is possible to know the pressure in both zones at any time  $n$ , or at least at the  $ns$  that takes to converge to a solution.
- The acoustic conditions, after a modification remain stationary or at least they vary sufficiently slowly [124].

In the PSZ literature, the use of iterative algorithms can be found in a few notable cases. For example, in 2006 Elliott *et al.* [3], using a similar system to the one studied in this dissertation, employed the FxLMS algorithm to calculate the filters that minimise the pressure at one ear of a dummy head located in an adjacent seat (the  $DZ$ ). As it was previously mentioned, in this dissertation the focus is put in controlling the resulting pressure in both zones, whilst in Ref. [3], the authors focus on the contrast maximisation. Nevertheless, the advantages of employing an adaptive algorithm (to compensate for small errors in the measured TFs) were already highlighted by the authors.

In 2017 Buerger, *et al.* [109] applied a modified version of the iterative frequency domain least-squares algorithm -previously used for room equalisation- in the context of multizone sound reproduction. Their method proposed the use of a weighted least-squares solution providing the flexibility to control the reproduction error both in the spatial and temporal domains. This formulation allowed to reduce perceptually undesirable pre-echoes without affecting the  $BZ$ . Additionally, in 2019 Schneider *et al.* [110] proposed to solve the PSZ problem with a novel algorithm computing update step in the frequency domain using an approximation with the exact determination of the error in the time-domain. This algorithm also included an adaptive regularisation and it was studied to generate one  $BZ$  and one  $DZ$  using three scenarios: two with measured IRs (16 loudspeaker linear array) in a room and one with simulated IRs (60 loudspeaker surrounding the zones).

Although the aforementioned studies propose to calculate the filter combination using non-deterministic approaches, the filters are still applied using a feed-forward control technique; *i.e.*, the iterative methods are used as a “better way” to calculate the filters, however, one falls in the same conflicts reviewed in Chapter 4, under a modification of the acoustic environment one needs to either re-calculate the filter combination or to have them stored.

Another example of implementation of iterative algorithms for PSZ are the works presented by Rousel [111] in 2019 where he studied experimentally the possibility to use different adaptive algorithms to different types of listener displacements. He showed that it was possible to use the LMS algorithm to adapt moving zones. However, in this study it was assumed that the zones moved in free-field and it was possible to know accurately the IRs at every moment.

Lastly, it is necessary to highlight that in 2017 Olsen and Møller [26] (and again Møller in 2019 [94]) mentioned the possibility to use a feed-back loop in order to accommodate to general changes in the sound field and very insightfully highlighted the challenge of the online-knowledge of the pressures in the zones. In addition, the possibility of their implementation for the generation of PSZ was also mentioned by Betlehem in 2015 [33], but ruled out due to the number of channels involved.

The use of iterative algorithms to adapt the filter coefficients of a feed-forward control can be found in other applications. For example, in the SFS domain, the use of adaptive algorithms to adapt the

coefficients of an open-loop system was studied by Gauthier and Berry in 2008 implementing the gradient descent iterative-algorithm [83] (from the basis of the work published in 2006 [77]). The goal of the methods proposed was to overcome the inherent free-field assumption of the WFS method which results in a reduction of the accuracy of the sound field reproduction in real situations using a closed-loop compensation. In these studies the authors propose a custom  $\mathbf{r}[n]$  that re-injects a scaled part of the WFS solution in each iteration.

In addition, the use of iterative algorithms was also studied by Cheer in 2012 in order to track the changes in the frequency and spatial distribution of the primary field that occur as the engine speed and load is varied in the control of engine noise [14]. Even more so, he already highlighted the physical similarities between active noise control (ANC) and audio reproduction [14].

Furthermore, it is possible to find several studies implementing these algorithms for ANC in headrests, for example in [159, 160, 161]. Moreover, in 2017 Antoñanzas *et al.* [168] presented a practical implementation of a personal ANC system to generate a local quiet zone on each seat close to the listener's ears in a two-seat configuration.

### The problem in hand

If one analyses the variations in a car cabin, some of these changes can be assumed to be considerably slow: temperature, loudspeaker wear, etc. Additionally, after some of the more drastic changes, the cabin can be seen as a stationary environment, *e.g.* as reviewed in Chapter 4, after the seat position is adjusted, they remain in that position. Hence, it is possible to use a recursive algorithm in order to re-calculate the filters after a modification [124].

In 1992, Nelson *et al.* described the theoretical basis for the design of adaptive digital filters used for the equalisation of the response of multichannel sound reproduction systems, which provides the possibility to design the filters *in-situ* [144] also known as the FxLMS. The main advantage of the FxLMS is, instead of requiring an accurate characterisation of the acoustic environment as in the deterministic approaches (also referred to as plant), it only needs an imperfect knowledge of the environment (a plant model) to calculate the filters. Therefore, the aim of the following development is to take advantage of this characteristic and, by only providing a plant model, calculate the filter combination under “unknown” acoustic-conditions.

Many other adaptive filter algorithms could be used to solve the problem in hand; *e.g.* other versions of the FxLMS algorithm (Modified FxLMS, adjoint LMS), the filtered-x recursive least squares (FxRLS), among many others [108]. Additionally, both methods proposed by Buerger [109] and Schneider [110] could also be considered to solve the invariant problem as in Chapter 3.

Although these other methods based on the FxRLS calculate the optimal filters in a least-squared sense as the PM, in this thesis the FxLMS algorithm is chosen, not only because it is considered as the benchmark to which most adaptive filtering algorithms are compared [108] but mainly because it requires considerably less computational complexity than the FxRLS <sup>2</sup> and it presents smaller sensitivity to unknown disturbances [124], both reasons make this algorithm more attractive for the car application.

One of the problems/limitations of the FxLMS is that it exhibits frequency dependent convergence behaviour [169]. Furthermore, it was observed in Ref. [144] that it is difficult for the algorithm to produce an effective equalisation at very low frequencies, which is exactly the frequency band of interest of our system (this problematic is re-addressed in detail below).

---

<sup>2</sup>The FxLMS does not require a matrix inversion, unlike for example the RLS algorithms that implement Woodbury's identity, also known as the matrix inversion lemma [124]

## 5.2 Adaptive PSZ: FxLMS theory

As mentioned above, the core of the adaptive algorithms, and where the distinctions among the different categories lay, is in the calculation of the update term  $\Upsilon[n]$ . In this chapter, one focuses on the FxLMS iterative algorithm. The following development is mainly based on the works of Nelson 1992 [144] and Bouchard 2000 [108].

To develop the FxLMS theory, one needs to bring back the expression of the pressure from Eq. 2.3

$$p_k[n] = \sum_{s=1}^S \sum_{l=1}^L \sum_{m=0}^{M-1} w_{s,l}[m] \sum_{i=0}^{I-1} h_{k,l}[i] x_s[n-m-i]. \quad (5.2)$$

Moreover, the  $(K \times 1)$  vector of the  $K$  sampled pressures can be written as

$$\mathbf{p}[n] = [p_1[n], p_2[n], \dots, p_K[n]]^T. \quad (5.3)$$

As in Chapter 2.3.3, the aim of the PSZ system studied in this thesis is to find the best combination of  $w_{s,l}[m]$ , to filter  $x_s[n]$  in order to minimise an error. Therefore, one needs to introduce the  $(K \times 1)$  vector of desired sampled pressures ( $\mathbf{d}[n]$ ) for the  $K$  microphones as

$$\mathbf{d}[n] = [d_1[n], d_2[n], \dots, d_K[n]]^T, \quad (5.4)$$

from Eq. 5.3 and Eq. 5.4, it is possible to define the estimation of the error as

$$\mathbf{e}[n] = \mathbf{d}[n] - \mathbf{p}[n]. \quad (5.5)$$

Using a matrix notation, one can express Eq. (5.2) as

$$p_k[n] = \mathbf{r}_k^T[n] \mathbf{w}[n] \quad (5.6)$$

being  $\mathbf{w}[n]$  the  $M$  coefficients time dependent  $(MSL \times 1)$  vector of concatenated FIR filters, arranged as

$$\begin{aligned} \mathbf{w}^T[n] = & \\ & [w_{1,1}[0], \dots, w_{1,1}[M-1], \dots, w_{1,L}[0], \dots, w_{1,L}[M-1], \\ & \dots, w_{S,1}[0], \dots, w_{S,1}[M-1], \dots, w_{S,L}[0], \dots, w_{S,L}[M-1]], \end{aligned} \quad (5.7)$$

and  $\mathbf{r}_k[n]$  the  $(MSL \times 1)$  vector of concatenated filtered signals expressed as

$$\mathbf{r}_k^T[n] = [\mathbf{r}_{k,1,1}^T[n] \ \dots \ \mathbf{r}_{k,L,1}^T[n] \ \dots \ \mathbf{r}_{k,1,S}^T[n] \ \dots \ \mathbf{r}_{k,L,S}^T[n]] \quad (5.8)$$

where  $\mathbf{r}_{k,l,s}[n]$  is the  $(M \times 1)$  vector expressed as

$$\mathbf{r}_{k,l,s}^T[n] = [r_{k,l,s}[n], r_{k,l,s}[n-1], \dots, r_{k,l,s}[n-M+1]], \quad (5.9)$$

being

$$r_{k,l,s}[n] = \sum_{i=0}^{I-1} h_{k,l}[i] x_s[n-i]. \quad (5.10)$$

Using Eqs. (5.3) and (5.6) it is possible to write the  $(K \times 1)$  vector of sampled pressures ( $\mathbf{p}[n]$ ) at time  $n$  over the  $K$  microphones as

$$\mathbf{p}[n] = \begin{bmatrix} \mathbf{r}_1^T[n] \\ \mathbf{r}_2^T[n] \\ \vdots \\ \mathbf{r}_K^T[n] \end{bmatrix} \mathbf{w}[n] \quad (5.11)$$

or more concisely as

$$\mathbf{p}[n] = \mathbf{R}[n] \mathbf{w}[n], \quad (5.12)$$

where  $\mathbf{R}$  is the  $K \times MSL$  matrix. Replacing Eq. 5.12 in Eq. 5.5

$$\mathbf{e}[n, \mathbf{w}] = \mathbf{d}[n] - \mathbf{R}[n] \mathbf{w}[n]. \quad (5.13)$$

Using Eq. 5.13, the cost function to minimise can be expressed as

$$\mathcal{J}[n, \mathbf{w}[n]] = |\mathbf{d}[n] - \mathbf{R}[n] \mathbf{w}[n]|^2. \quad (5.14)$$

The steepest-descent method can be used to solve the quadratic form in Eq. (5.14) by iteratively updating the coefficients of  $\mathbf{w}[n]$  [164]. Therefore, the update equation of the multichannel FxLMS is

$$\mathbf{w}[n + 1] = \mathbf{w}[n] + \mu \mathbf{R}^T[n] \mathbf{e}[n] \quad (5.15)$$

being  $\mu$  a scalar also known as the step-size parameter or convergence gain and  $\mathbf{e}[n]$  given in Eq.5.5.

One might wonder how is this system going to adapt to unknown acoustic conditions if the matrix  $\mathbf{R}$  is built with the system's IRs (Eq. 5.11). However, Nelson *et al.* already mentioned that to solve Eq. (5.15) it is only necessary to have an imperfect knowledge of  $\mathbf{R}$  [144]. Therefore, Eq. (5.15) becomes

$$\mathbf{w}[n + 1] = \mathbf{w}[n] + \mu \tilde{\mathbf{R}}^T[n] \mathbf{e}[n], \quad (5.16)$$

where the ( $\tilde{\phantom{x}}$ ) establishes that is only a model/approximation of the plant [144].

Therefore, the update term  $\Upsilon = \mu \tilde{\mathbf{R}}^T[n] \mathbf{e}[n]$ , establishes that it is necessary, to calculate for each iteration/sample  $n$  the error ( $\mathbf{e}[n]$ ) and the filtered-x matrix  $\tilde{\mathbf{R}}$ .

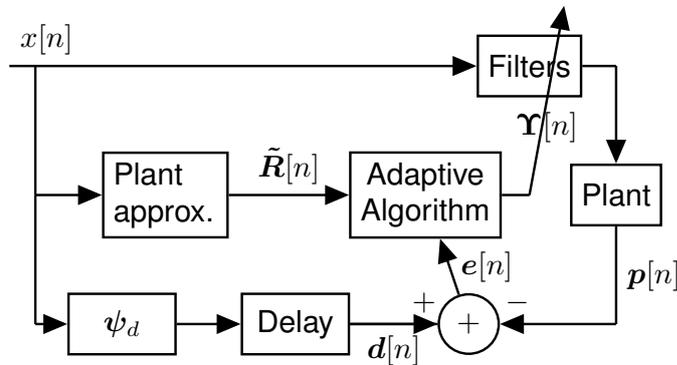


Figure 5.1: FxLMS block diagram

In Fig. 5.1 is presented the block diagram of the system, where it is possible to identify that the PSZ corresponds to the class of applications under the name of inverse modelling [124]. Moreover, it is shown the fact that the system requires to have a real-time knowledge of  $\mathbf{p}[n]$  and also that the versions of the input are filtered by the plant model to obtain  $\tilde{\mathbf{R}}[n]$ . In addition, it is also necessary to define  $\mathbf{d}[n]$ , for this, a new parameter is introduced: the desired filter vector ( $\psi_d$ ) which corresponds to  $K$  filters applied to  $x_s[n]$ .

It is necessary to highlight that the method is valid as long as  $S \leq L \leq K$ , *i.e.* an overdetermined system. Additionally, if one wanted to reproduce one different program on each zone ( $S = 2$ ), the FxLMS's formulation allows to solve both problems simultaneously; or in PSZ terms  $Z_1 = BZ(x_1)$ ,  $Z_2 = DZ(x_1)$  and  $Z_1 = DZ(x_2)$ ,  $Z_2 = BZ(x_2)$ . Nevertheless, to be consistent with the PSZ literature -and the previous chapters- only the first case is reviewed; *i.e.*  $S = 1$ ,  $Z_1 = BZ(x_1)$  and  $Z_2 = DZ(x_1)$ . Furthermore, the filters to calculate the resulting metrics from Sec. 2.3.2 are assumed to be the ones corresponding to the last iteration of the input signal.

To sum up, one needs:

- To calculate  $\tilde{\mathbf{R}}[n]$ : the filtered-x matrix using the plant approximation.
- To calculate  $\mathbf{d}[n]$ : The input ( $x[n]$ ) filtered by the desired filter vector ( $\psi_d$ ).
- The step size parameter ( $\mu$ ).
- The initial coefficients ( $\mathbf{w}[0]$ ).

Therefore the effects of these definitions are studied using the IRs acquired in the car cabin (from Sec.4.2).

### 5.3 FxLMS definitions

The aim of this section is to find the best combination of the aforementioned parameters to fulfil the original requirements (see Chapter 1).

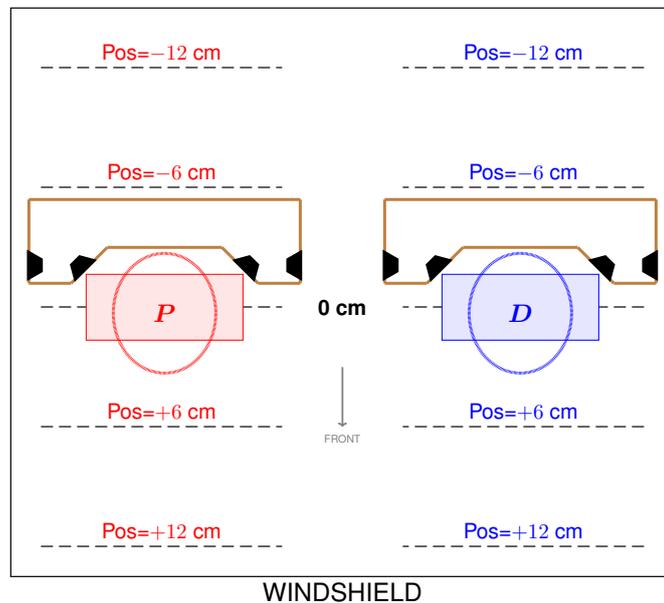


Figure 5.2: Schematics: changes in the seat positions (copy of Fig. 4.1a).

Moreover, the following results are calculated using measured IRs in a car cabin either at one position (Pos = 0), to study the effects of the step-size parameter and the convergence of the algorithm, or using the five seat-symmetric movements in order to evaluate the changes in the acoustic environment (see Fig.5.2). Nevertheless, some conclusions throughout this section are supported by using simulated IRs detailed in Appendix F.

In this first study, it is also assumed that 36 microphones are available in the cabin as for the previous chapters (See Fig. 4.11), this strong assumption is re-addressed in Sec. 5.4.

#### 5.3.1 FxLMS: step size parameter and convergence characteristics

In order to study the convergence characteristics of the FxLMS and to illustrate the effects of  $\mu$ , the FxLMS is used as the algorithm to calculate the filter combination at one position (Pos = 0 in Fig. 5.2), with the headrest in the bottom position and without any passengers on the rear seats). In other words, the FxLMS is used as an alternative to the deterministic approaches to calculate the optimal filter combinations.

As an attempt to solely observe the effect of  $\mu$  in this analysis  $\tilde{\mathbf{R}} = \mathbf{R}$ , *i.e.* the inaccuracies of the plant are removed. Additionally, it is assumed that there is no-prior knowledge about the initial filters  $\mathbf{w}[0] = \mathbf{0}$  and the signal is defined as a 10 s random noise.

The step-size parameter, also known as the convergence coefficient [167], plays a critical role in assuring convergence of the algorithm, or failing to do so [124]. It has been shown (*e.g.* refs. [124, 167]) that the algorithm will converge only if

$$0 < \mu < \frac{2}{T_{max}}, \quad (5.17)$$

being  $T_{max}$  the largest eigenvalue of the autocorrelation matrix  $\mathbb{E} \left\{ \tilde{\mathbf{R}}^T[n] \mathbf{R}[n] \right\}$  being  $\mathbb{E} \{ \cdot \}$  the mathematical expectation. Therefore, using Eq. 5.17 it is possible to calculate the convergence condition. For the studied case, it resulted in  $\mu < 8.7 \times 10^{-5}$ , however, it was later observed empirically that for the algorithm to be stable  $\mu \leq 6 \times 10^{-5}$ , hence, only these values are analysed.

Last but not least, the desired filter vector is defined as in the deterministic methods. To achieve a mono-signal in the *BZ*,  $\psi_{dBZ}$  is defined as the resulting IR when the two loudspeakers closer to the zone are driven in phase ( $P_R$  and  $P_L$ ). In the *DZ*,  $\psi_{dDZ} = \mathbf{0}$ . Nevertheless, the definition of  $\psi$  is studied in detail later. Lastly, the delay (see Fig.5.1) did not compromise the performance significantly, therefore, it was fixed to 20 samples to match the one applied in the deterministic approach.

All the definitions for this preliminary study are summarised in Table 5.1.

<b>FxLMS parameters: preliminary study</b>	
Sampling Frequency	2450 Hz
Filter Coefficients	M = 200
Delay	20 samples
CP	Car cabin: seats at Pos = 0 cm
EP	Car cabin: seats at Pos = 0 cm
$\psi_{dBZ}$ filter (BZ)	From each EP ( $P_L$ and $P_R$ driven in phase )
$\psi_{dDZ}$ filter (DZ)	Vector of zeros ( $\mathbf{0}$ )
Input	10 s random signal
Filters at $n = 0$	$\mathbf{w}[0] = \mathbf{0}$
Plant approx.	Ideal plant ( $\tilde{\mathbf{R}} = \mathbf{R}$ )
Step-size param.	$\mu = [6 \times 10^{-5} - 1 \times 10^{-5}]$
Mics. available	$K = 36$
Mics. evaluation	$K = 36$

Table 5.1: Summary of the FxLMS parameters for the study of the  $\mu$  effects.

One proceeded to plot the so-called learning curve which is defined as the modulus of the error squared as a function of  $n$ , in other words, the resulting value of the cost-function in Eq. 5.15 for the six different  $\mu$  parameters smoothed with a moving average filter to improve readability. To have a reference to the problem in hand, although the FxLMS when  $n \rightarrow \infty$  approaches asymptotically the value of the Weiner solution (which is optimal in a mean-square sense) [124], the modulus of the error squared obtained with the BACCPM algorithm is showed as a reference (which is optimal in a least-squares sense).

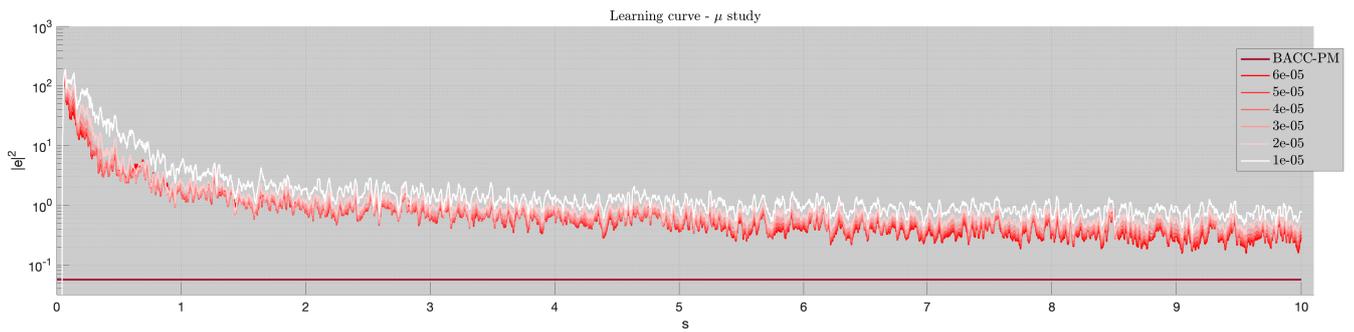


Figure 5.3: Learning curve for different  $\mu$  values at Pos = 0.

Firstly, it is possible to observe in Fig. 5.3 the typical exponential decay of the stochastic gradient descent methods [124]. Although the use of the small  $\mu$  ensures the convergence of the FxLMS algorithm [124], the larger  $\mu$  the smaller the resulting error, *i.e.* the faster the algorithm approaches the optimal filter combination. Nevertheless, in none of the studied  $\mu$ s, it is possible to achieve -after 10 s- the resulting reproduction error obtained with the BACCPM.

In order to extend these conclusions to the PSZ literature, one proceeded to plot both the resulting AC and  $E_{rr}$  using the last filter combination obtained with the FxLMS, *i.e.* after 10 s.

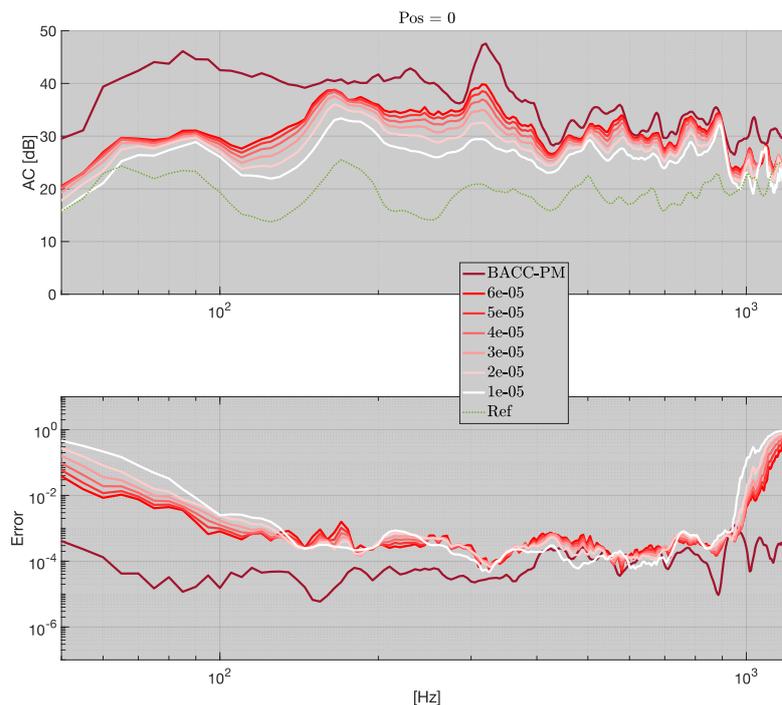


Figure 5.4: Results: Metrics at Pos = 0 using FxLMS after 10 s.

The conclusions achieved from the learning curves in Fig. 5.3 can be extended to the studied PSZ-metrics in Fig. 5.4. Firstly, none of the studied  $\mu$ s were able to match the levels neither of AC nor  $E_{rr}$  achieved with the BACCPM, especially not below  $\approx 150$  Hz where the response of the loudspeaker FRF starts dropping above 900 Hz where the downsampling filter is also decreasing the level. To illustrate this, four FRF are shown in Fig. 5.5 of the same relative-microphone in each of the four ears of the passengers when only one loudspeaker ( $P_R$ ) is driven without any processing.

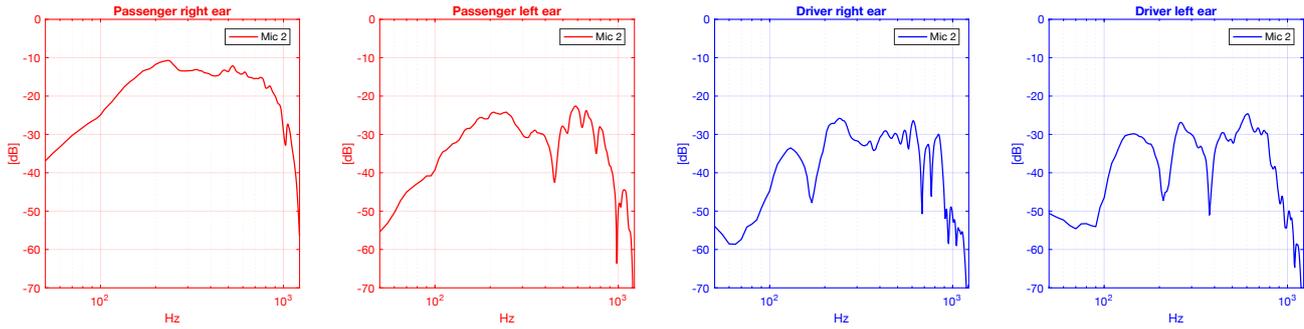


Figure 5.5: FRF when only driving loudspeaker  $P_R$  without any processing for the four microphones located at the ears, (microphones 2 in Fig. 4.11).

The frequency-dependent behaviour of the FxLMS in this application can be attributed to the band-pass nature of the all the system FRFs (see Fig. 5.5). In addition, in Appendix F it is possible to find the same analysis using perfectly-flat free-field monopole simulations where the frequency dependence is almost non-existent compared to the same simulated monopoles but with a second-order high-pass filter applied where similar results to the ones obtained in Fig. 5.4 are found.

Regarding the  $E_{rr}$ , in Fig. 5.4 the behaviour is even more frequency-dependent. For the lowest part of the spectrum (due to the loudspeaker’s FRF) it is possible to observe that there is a drop in performance and that the value of the step-size plays a major role. Additionally, a similar behaviour can be observed above 1 kHz -in this case due to the down-sampling filter. However, between 200 Hz and 900 Hz the value of the step size is not affecting much the reproduction error.

Therefore, it is necessary to use a step-size parameter small enough to avoid divergence but large enough that it can achieve the expected results, especially on the extremes of the studied frequency spectrum. To make sure that the system remains stable, the same analysis was done using a 300 s random signal.

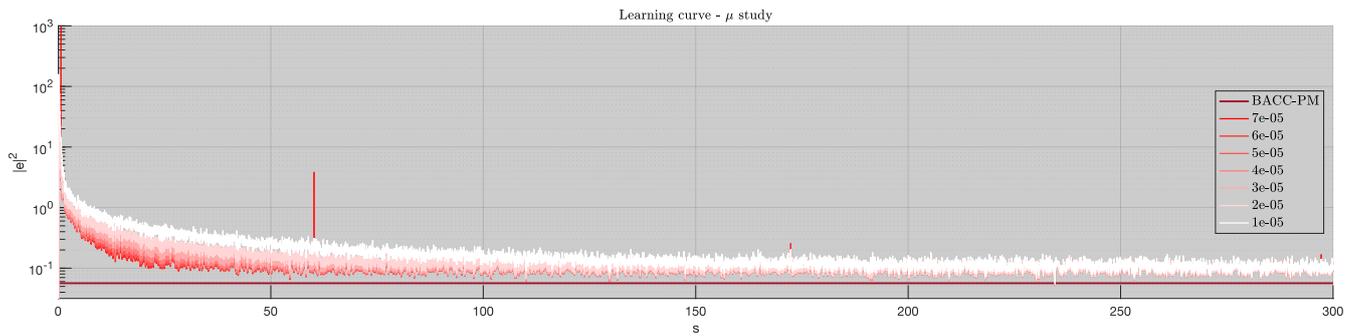


Figure 5.6: Learning curve for different  $\mu$  values at Pos = 0, 300 s long.

Firstly, it is possible to notice in Fig. 5.6 two “spikes” -one after 60 s and one after 172 s- which correspond to  $\mu = 6 \times 10^{-5}$ . Although the algorithm is capable of adjusting to these uncertainties, in order to avoid any possible instability  $\mu = 5 \times 10^{-5}$  is chosen as the best compromise. Secondly, when comparing among the different step size parameters, the larger values achieve the lower levels of  $|e|^2$  faster, while the smaller  $\mu$  take considerably longer to achieve the same values.

Moreover, the longer the number of iterations (longer input signal), the smaller the lower levels of  $|e|^2$  the system achieves. In addition, one proceeded to study the resulting metrics after a 300 s signal for the different values of  $\mu$ .

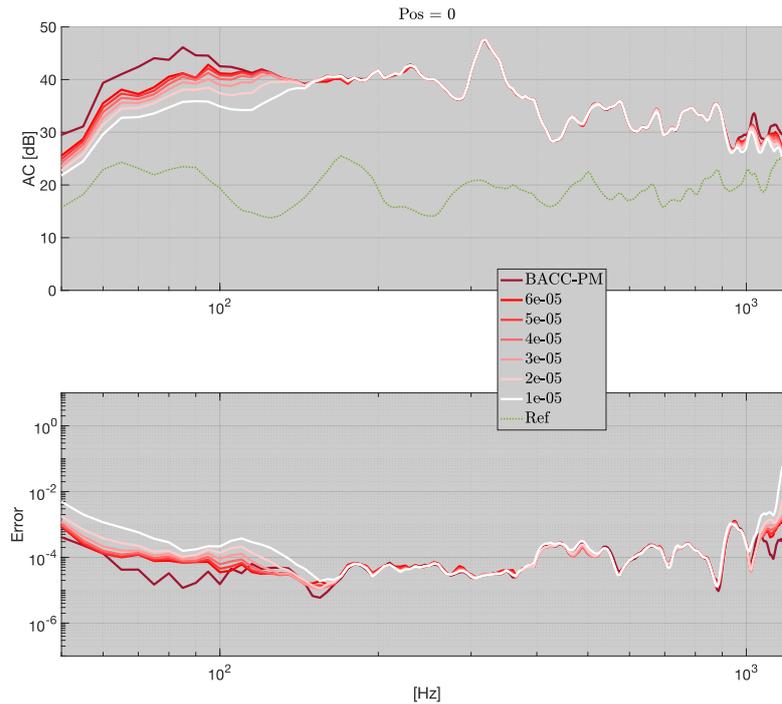


Figure 5.7: Results: Metrics at Pos = 0 using FxLMS after 300 s.

Even though in Fig. 5.6 it seems that all the values of  $\mu$  arrived to the same -or similar-  $|e|^2$ , when analysing the studied metrics (Fig. 5.7) there is still a noticeable difference in the lowest part of the spectrum (below  $\approx 150$  Hz) where still the step-size parameter plays a mayor role on the performance, both in the resulting  $AC$  and  $E_{rr}$ .

To sum up, a preliminary study was made using measured IRs in order to study the influence of the step-size parameter and the convergence behaviour of the FxLMS. The largest possible  $\mu$  was found and studied using real measured data to calculate the PSZ filters “from-scratch” ( $\mathbf{w}[0] = \mathbf{\emptyset}$ ).

One of the limitations of the FxLMS -highlighted in Fig. 5.4- is that this algorithm exhibits a frequency dependent convergence behaviour which shows a degradation in the overall performance [144, 169], particularly worrying for the application in hand below 150 Hz. Additionally, using an extremely long signal it was observed that this lack of performance in the lowest part of the spectrum was still present, especially for the smaller values of  $\mu$  where the algorithm is more stable.

It is possible to find numerous publications studying the choice of the step-size parameter and how to improve the convergence and stability of the FxLMS. For example, Elliott in Ref. [167] proposed to add an effort constraint to the cost function in Eq. 5.14. Other authors (*e.g.* Refs. [170, 171]) propose to use a variable step-size parameter which can result in faster-convergence, better performance, smaller sensitivity to noise and smaller sensitivity with respect to the system parameters. Additionally, in Ref. [169], the authors proposed to include an eigenvalue equalisation to improve the frequency dependence of the algorithm. Although these approaches are out of the scope of this thesis, the author does not rule out their use in a complementary way in future works.

Lastly, in this preliminary study the FxLMS algorithm was used as an alternative to calculate the filter combinations “from-scratch”. However, it does not present any noticeable advantage over the BACCPM other than reducing the computational cost and not having to invert any matrix, especially because to achieve comparable performance it takes extremely long time. Moreover, if one already knows the plant  $\tilde{\mathbf{R}} = \mathbf{R}$ , it is possible to simply use the deterministic approaches as reviewed in Chapter 4 or the iterative methods proposed in [109, 110] to obtain the filter combination.

However, as it has been pointed out, the goal of using the FxLMS is to obtain the filter combination when one does not have access to the plant, *i.e.*  $\tilde{\mathbf{R}} \neq \mathbf{R}$ .

### 5.3.2 FxLMS: the initial coefficients and the plant model

Before explaining the method proposed to improve the frequency-dependent convergence of the FxLMS it is necessary to bring back the notion of calibration position (CP) and evaluation position (EP) defined in Sec. 4.2.

In agreement with Chapter 4, one is going to assume that it is possible, for one CP to use a deterministic approach, comprising: the acquisition of all the necessary impulse responses and the calculation of the optimal filters (from this point forward referred to as calibration filters or  $w_{CP}$ ).

As it was reviewed in the previous section, the main reason of using the FxLMS algorithm is to take advantage of the fact that it is only necessary to have an imperfect knowledge of the plant. Therefore, in this thesis it is proposed to use as a plant model the IRs acquired at the CP, *i.e.* the ones used to calculate the calibration filters. It is worth mentioning that the use of the plant simulated in free-field conditions was also considered as a plant model but the algorithm became too unstable, hence its use is ignored from this study.

Moreover, it is important to notice that as long as the acoustic conditions do not change -stays at CP- the calibration filters are optimal in the least-squares sense [124]. However, in the presence of a change in the acoustic conditions (at EP) the FxLMS algorithm is used, not to re-calculate the optimal filter combinations from scratch as done when studying  $\mu$ , but to calculate the new filter combination using as a starting point the calibration filters obtained at CP.

For this report, the CP is taken when both seats are located in the centre of the rails (see Fig. 5.2), with the headrest in the lowest position and without any passengers on the rear seats. Moreover, it is assumed that as only one filter combination is calculated -one CP-, it is possible to afford to calculate the filters using the time domain formulation explained in Sec. 2.3.3.3, the BACCPM. In addition, the step-size parameter used is constant for every iteration and is chosen empirically trying to obtain a  $\mu$  as large as possible without compromising the stability of the system. In addition, from this point forward it is arbitrarily assumed that after 10 s the system should achieve a comparable performance to the one achieved with the BACCPM.

In order to illustrate the contribution of the initial coefficients, the system is evaluated using the five symmetric movements reviewed in Sec. 4.3. Additionally, it was observed in Chapter 3 that the filters close to the BZ were always almost perfect  $\delta[n]$  and in Sec. 4.6 that it was possible to remove them from the calculation, therefore the use of two theoretical  $\delta[n]$  as initial coefficients ( $w[0]$ ) at  $P_R$  and  $P_L$  is also analysed. All the studied definitions are summarised in Table 5.2.

FxLMS parameters: $w[0]$ study	
Sampling Frequency	2450 Hz
Filter Coefficients	$M = 200$
Delay	20 samples
CP	Car cabin: seats at Pos = 0 cm
EP	Car cabin: five symmetric seat positions
$\psi_{dBZ}$ filter (BZ)	From each EP ( $P_L$ and $P_R$ driven in phase )
$\psi_{dDZ}$ filter (DZ)	Vector of zeros ( $\emptyset$ )
Input	10 s random signal
$w[0]$	$\emptyset$ vs. $2 \times \delta[n]$ vs. $w_{CP}$
Plant approx.	IRs from CP
Step-size param.	$\mu = 5 \times 10^{-5}$
Mics. available	$K = 36$
Mics. evaluation	$K = 36$

Table 5.2: Summary of the FxLMS parameters for the study of the initial coefficients.

One proceeded to plot the resulting  $AC$  and  $E_{rr}$  after 10 s for the FxLMS. In order to have a reference of behaviour, the results from Sec. 4.3.1 are showed in Fig. 5.8, comprising the Re-calibrated approach -the ideal case when it is possible to re-acquire all the IRs- and the Fixed approach -when  $\mathbf{w}_{CP}$  are used at every position- defining the shadowed area where the system “should perform” to provide an improvement.

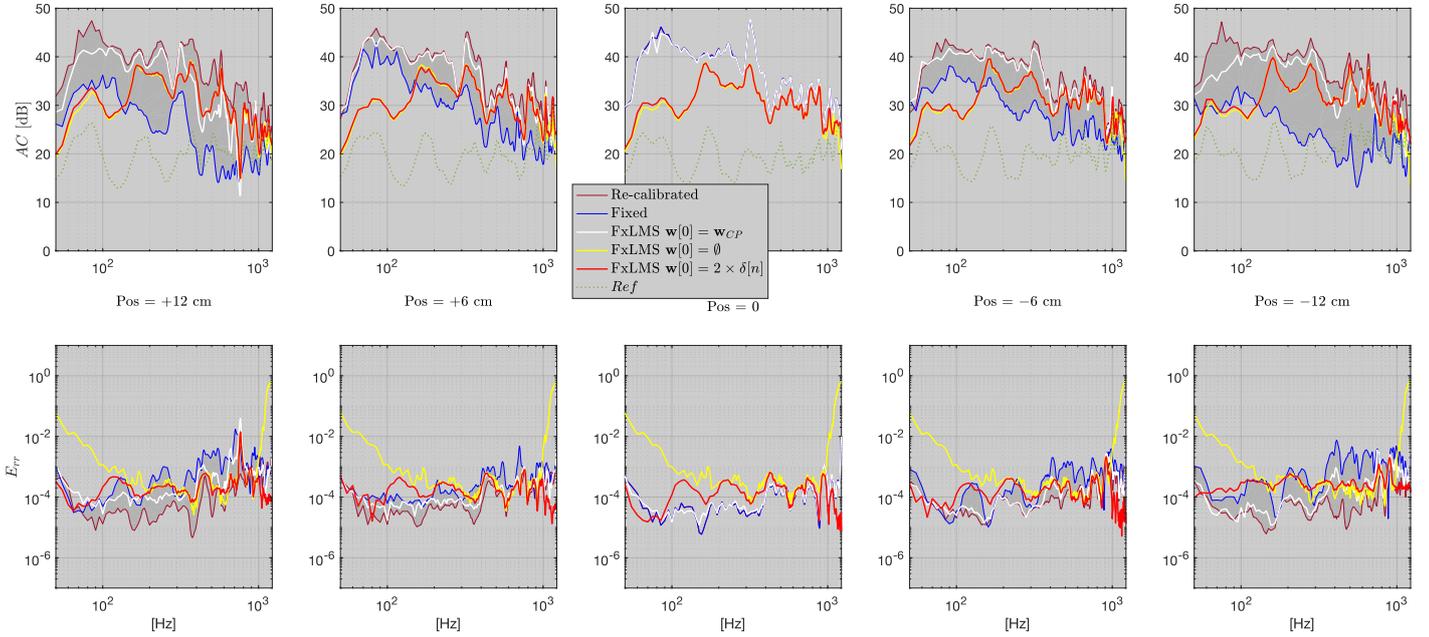


Figure 5.8: FxLMS results depending on the definition of the initial coefficients. Resulting  $AC$  and  $E_{rr}$  for symmetric seat positions.

Firstly, it is possible to observe in Fig. 5.8 a noticeable improvement when  $\mathbf{w}[0] = \mathbf{w}_{CP}$  compared to the other two initial coefficients definitions. Both the  $AC$  as well as the  $E_{rr}$  is considerably improved especially below 300 Hz regardless of the seat-position. Therefore, the proposed method of feeding the calibration filters showed to be effective to improve the frequency-dependent convergence of the FxLMS.

Moreover, when  $\mathbf{w}[0] = \mathbf{w}_{CP}$ , the FxLMS is capable of improving the Fixed approach (white curves within the shadow area in Fig. 5.8) in the four unknown studied acoustic conditions, especially at Pos =  $\pm 6$  cm where results are almost identical to the Re-calibrated ones.

However, at the extreme positions ( $\pm 12$  cm) when  $\mathbf{w}[0] = \mathbf{w}_{CP}$  results in some noticeable  $AC$  differences, especially in Pos +12 cm around 450 Hz and around 750 Hz and in Pos = -12 cm around 450 Hz. These differences are indication of the fact that the “imperfect knowledge” of the plant model can be “too imperfect”, *i.e.* at these frequencies, the plant model is too far from the acoustic conditions at EP and it cannot properly calculate the filters. Additionally, one can observe that at the lowest frequencies the algorithm is not achieving  $AC$  as good as the re-calibrated, but its resulting levels still considerably outperform the fixed ones.

Regarding the  $E_{rr}$ , in spite of the fact that its deterioration is not as pronounced as the  $AC$  (re-calibrated approach is close to the fixed approach), the FxLMS when  $\mathbf{w}[0] = \mathbf{w}_{CP}$  is capable of maintaining the order of magnitude in all the positions and improve over the fixed approach.

On the other hand, when only two  $\delta[n]$  were used as initial coefficients (red curves in Fig. 5.8) there is a considerable improvement in the  $E_{rr}$  in comparison to calculating the filters from scratch, especially on the ends of the frequency band. This is a direct consequence of the fact that the reproduction filters, in the way the problem is posed always result in almost-perfect  $\delta[n]$  function. However, regarding the  $AC$ , it achieves almost the identical levels to the  $\mathbf{w}[0] = \emptyset$ .

It is necessary to make a particular analysis to what happens at Pos = 0. When  $\mathbf{w}[0] = \emptyset$ , it was already studied that the FxLMS algorithm was not capable of achieving a comparable performance after

10 s (see Sec.5.3.1). On the other hand, when  $\mathbf{w}[0] = \mathbf{w}_{CP}$ , it was stated above that the calibration filters at this position are optimal in a least-squares sense and as the acoustic conditions do not change the FxLMS should not “move”, however, it is possible to observe a slight modification both in the  $AC$  as well-as the  $E_{rr}$ . Although these changes are negligible, their justification could be a combination of the following reasons: firstly because of the fact that the stochastic gradient descent (...) *once it reaches the local neighbourhood of the optimum solution, it continually wanders around that solution in a random-walk fashion and therefore never settles down to an equilibrium point* [124]. In addition, it could be also possible because the BACCPM is optimal in a least-squares sense whilst the FxLMS approaches the Wiener solution, when  $n \rightarrow \infty$ , which is optimal in the mean-square-error sense and might slightly differ.

In this analysis, it was showed that it is possible to considerably improve the frequency-dependent convergence nature of the FxLMS by feeding the initial coefficients from a deterministic approach algorithm (the BACCPM). The algorithm was capable of considerably improving the filter combinations after 10 s in spite of the fact that the acoustic conditions at these new EP were unknown. However, there is one key definition that has been overseen in this analysis that needs to be studied in detail, the desired filter vector.

### 5.3.3 FxLMS: desired filter vector

One of the original constraints of the system is to achieve a transparent system (see Chapter1), hence, when using the deterministic approach the desired pressure was defined for the  $BZ$  as the resulting pressure field of the sources the closest to the zone when driven in phase (see Sec. 3.1) and zeros ( $\mathbf{0}$ ) in the  $DZ$  to achieve a potential energy minimisation (see Sec. 2.3.3.2). However, defining  $\boldsymbol{\psi}_d$  to calculate the FxLMS becomes slightly more complicated.

On the one hand, the definition of desired filter vector in the  $DZ$  does not present any particular problem, one can take the desired definition from the deterministic approach (Sec. 2.3.3.2) and define  $\boldsymbol{\psi}_{dDZ} = \mathbf{p}_{dDZ} = \mathbf{0}$  resulting in the potential energy minimisation as in the deterministic approaches.

On the other hand, some explanations need to be given for the  $BZ$ . Firstly, it is important to notice that whilst in the deterministic approach algorithms one was defining a desired pressure  $\mathbf{p}_{dBZ}$  to calculate the error (see Sec. 2.3.3.2),  $\boldsymbol{\psi}_{dBZ}$  is the filter that is going to be applied to the input signal  $x[n]$  to obtain the desired pressure. Nevertheless, as in the deterministic approaches it is assumed a  $\delta[n]$  as input signal it is possible, in spite of the unit mismatching, to define  $\boldsymbol{\psi}_{dBZ} = \mathbf{p}_{dBZ}$ .

Moreover, in order to achieve a proper transparent system  $\mathbf{p}_{dBZ}$  was defined as the two sources the closest to the  $BZ$  when driven in phase (see Sec. 3.1), but to be perfectly accurate,  $\mathbf{p}_{dBZ}$  should be acquired at every new EP (as done in the re-calibrated approaches) because it differs from the one obtained at CP. To illustrate this difference, in Fig. 5.9 the amplitude response of the  $\mathbf{p}_{dBZ}$  is showed for six different microphones (three on each ear) for three different positions.

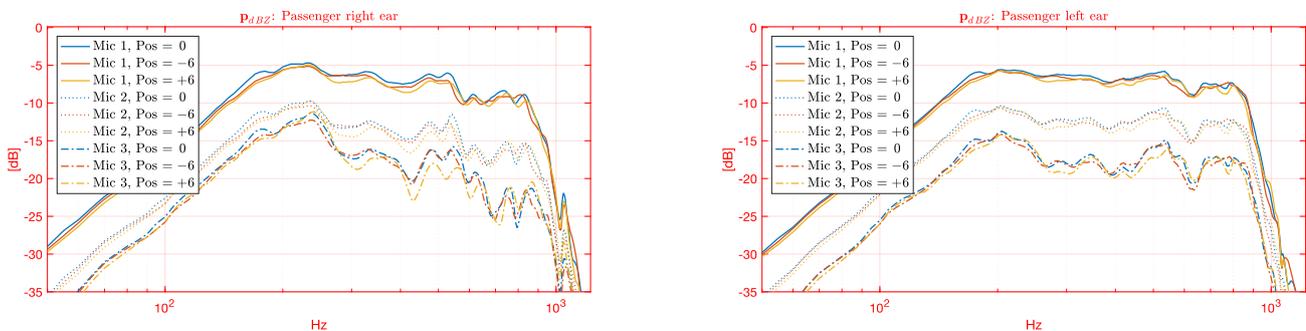


Figure 5.9: Desired pressures in the  $BZ$  ( $\mathbf{p}_{dBZ}$ ) for microphones 1 (bold), 2 (dotted) and 3 (dashed-dotted) (see Fig. 4.11) at three different seat positions Pos = 0 (blue), Pos = -6 (red) and Pos = +6 (yellow).

In Fig. 5.9, one can observe that the resulting  $\mathbf{p}_{dBZ}$  for three different microphones present amplitude

differences in all three microphones, in both ears, depending on the position of the seats.

Due to the fact that the microphones are placed in such proximity of the sources these differences are small, but the order of magnitude of  $E_{rr}$  handled throughout this thesis ( $\simeq 10^{-4}$ ) is also small. When using the FxLMS, these small differences affect the resulting  $E_{rr}$  because the algorithm is attempting an error minimisation in the  $BZ$  for a  $\mathbf{d}[n]$  that does not correspond to the EP.

FxLMS parameters: $\psi_{dBZ}$ study	
Sampling Frequency	2450 Hz
Filter Coefficients	$M = 200$
Delay	20 samples
CP	Car cabin: seats at Pos = 0 cm
EP	Car cabin: five symmetric seat positions
$\psi_{dBZ}$ filter (BZ)	$\psi_{dBZ}$ from: CP vs. from each EP ( $P_L$ and $P_R$ driven in phase )
$\psi_{dDZ}$ filter (DZ)	Vector of zeros ( $\mathbf{0}$ )
Input	10 s random signal
$\mathbf{w}[0]$	$\mathbf{w}_{CP}$
Plant approx.	IRs from CP
Step-size param.	$\mu = 5 \times 10^{-5}$
Mics. available	$K = 36$
Mics. evaluation	$K = 36$

Table 5.3: Summary of the FxLMS parameters for the study of the initial coefficients

In order to show this subtlety, two different definitions of  $\psi_{dBZ}$  are considered, one taking  $\mathbf{p}_{dBZ}$  at CP for every EP and one taking a dedicated  $\mathbf{p}_{dBZ}$  for each EP. A summary of the definitions used for this analysis can be found in Table 5.3.

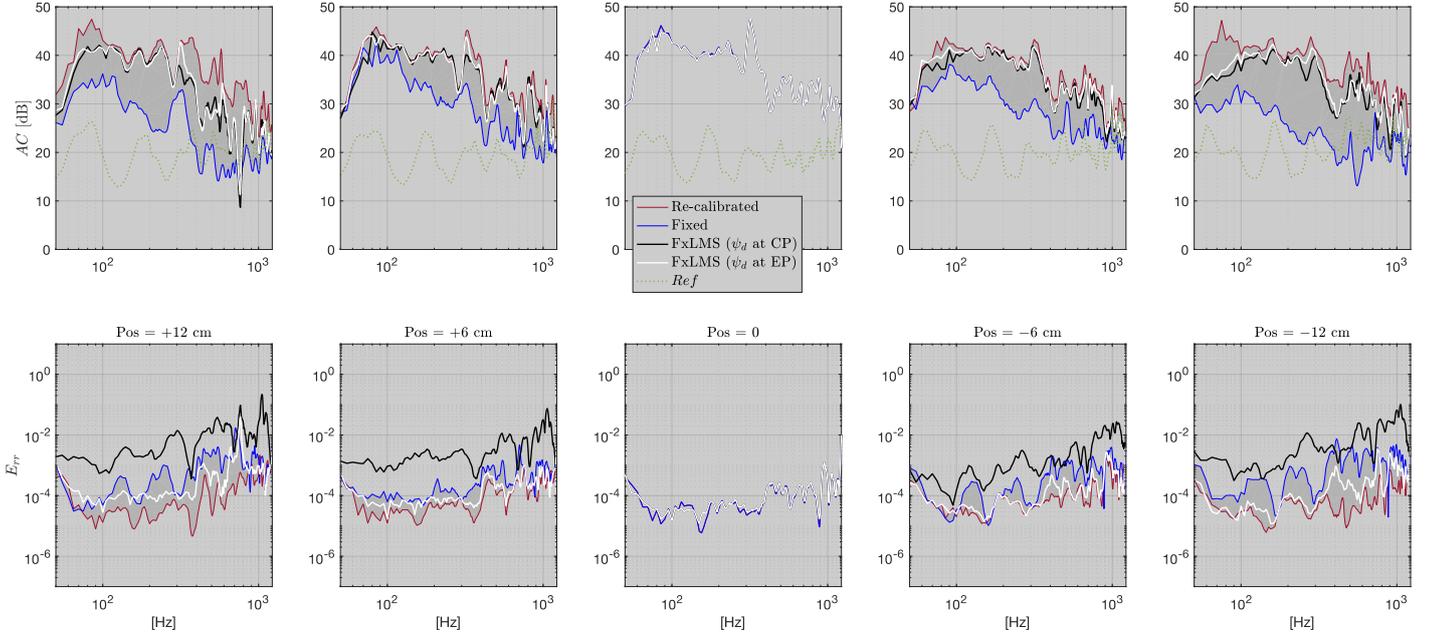


Figure 5.10: FxLMS results depending on the definition of the desired filters. Resulting  $AC$  and  $E_{rr}$  for symmetric seat positions.

Regarding the  $AC$ , it is possible to observe in Fig. 5.10 that regardless of the  $\psi_{dBZ}$  the algorithm

achieves almost the same levels for all the studied positions. However, when analysing the  $E_{rr}$  there is an almost homogeneous deterioration in all the positions surpassing the Fixed approach. This is due to the fact that the algorithm is minimising the  $E_{rr}$  of the  $\mathbf{d}[n]$  of a different acoustic condition.

In the author's opinion, it is not possible to state if the deterioration from taking one fixed  $\psi_{dBZ}$  (Fig. 5.10) alters the system's transparency. In order to make this decision, it would be necessary to assist the design with a series of psychoacoustic tests to evaluate if these differences are noticeable or not. Nevertheless, for the following studies it is assumed that a dedicated  $\psi_{dBZ}$  is defined for each EP.

## 5.4 FxLMS variable acoustics: microphone reduction

In the previous sections, all the FxLMS definitions were studied and the system was evaluated for five different seat-positions in a car cabin proving to be able to achieve a dedicated filter combination capable of maintaining the system's performance having 36 microphones in the cabin. However, as it was stated in Sec. 4.5, it is not realistic to assume that one can have this microphone outline permanently installed in the car cabin (see Fig. 5.11).

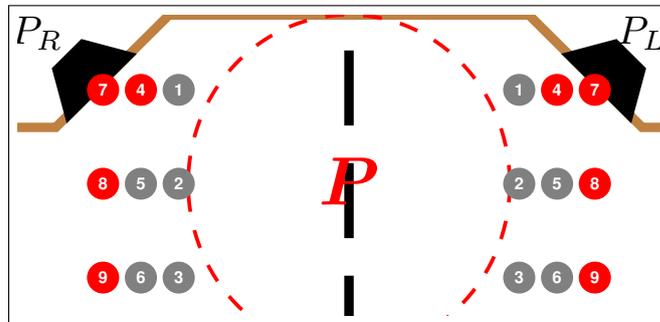


Figure 5.11: Schematic: microphone outlines passenger seat (copy of Fig.4.11)

Therefore, one is going to make a study to reduce the number of microphones in a car cabin. Following the results obtained in Sec. 4.5.1, the microphones contributing the most to the overall problem are microphones 7 and 9, additionally, it was observed that the addition of microphone 4 was also beneficial to the problem, therefore one is going to analyse the microphone outlines given in Table 5.4.

Id	$K$	Mic. outline	$\mu$
a)	36	1 to 9	$5 \times 10^{-5}$
b)	16	4, 7, 8 and 9	$8 \times 10^{-5}$
c)	12	4, 7 and 9	$8.5 \times 10^{-5}$
d)	8	7 and 9	$1 \times 10^{-4}$
e)	8	4 and 9	$1 \times 10^{-4}$

Table 5.4: FxLMS microphone outlines with corresponding step-size parameters.

The fact that the number of microphones used for the adaptation is changing, means that both the size of  $\tilde{\mathbf{R}}$  and the size of  $\mathbf{e}[n]$  are changing as well (see Eq. 5.16). Therefore it is necessary to change the step-size parameter accordingly, as for the previous sections, the value of  $\mu$  is also chosen empirically to the largest value possible that allows the five positions to converge to a solution, but a different value is used for each microphone combination. Although from outlines a) to e) the smaller the number of microphones the larger the  $\mu$ , this does not mean that the number of microphones is the only parameter defining the possible  $\mu$ .

The outlines chosen are the extension of the results obtained in Sec. 4.5.1, nevertheless, the other combinations gave acceptable results as well. All the parameters used for the calculations are listed in Table 5.5.

FxLMS parameters: microphone removal study	
Sampling Frequency	2450 Hz
Filter Coefficients	$M = 200$
Delay	20 samples
CP	Car cabin: seats at Pos = 0 cm
EP	Car cabin: five symmetric seat positions
$\psi_{dBZ}$ filter (BZ)	From each EP ( $P_L$ and $P_R$ driven in phase )
$\psi_{dDZ}$ filter (DZ)	Vector of zeros ( $\mathbf{0}$ )
Input	10 s random signal
$\mathbf{w}[0]$	$\mathbf{w}_{CP}$
Plant approx.	IRs from CP
Step-size param.	see Table 5.4
Mics. available	see Table 5.4
Mics. evaluation	$K = 36$

Table 5.5: Summary of the FxLMS parameters for the microphone removal study.

Lastly, as done in Sec. 4.5.1, one is going to assist the analysis by using the already defined  $AC_G$ . In order to solely evaluate the deterioration of reducing the number of microphones, the reference is chosen as the FxLMS results of configuration *a*) (from Table 5.4). It is important to mention that, regardless the combination used for the calculation, both the  $AC_G$  and the  $E_{rr}$  are always evaluated over the 36 original microphone outline (Fig. 5.11). Moreover, to evaluate the deterioration of the microphone-removal, outline *a*) in Table 5.4 is used as a reference to calculate the  $AC_G$ .

One proceeded to calculate the resulting  $AC_G$  and  $E_{rr}$  for five different seat-positions and for the different microphone outlines listed in Table 5.4.

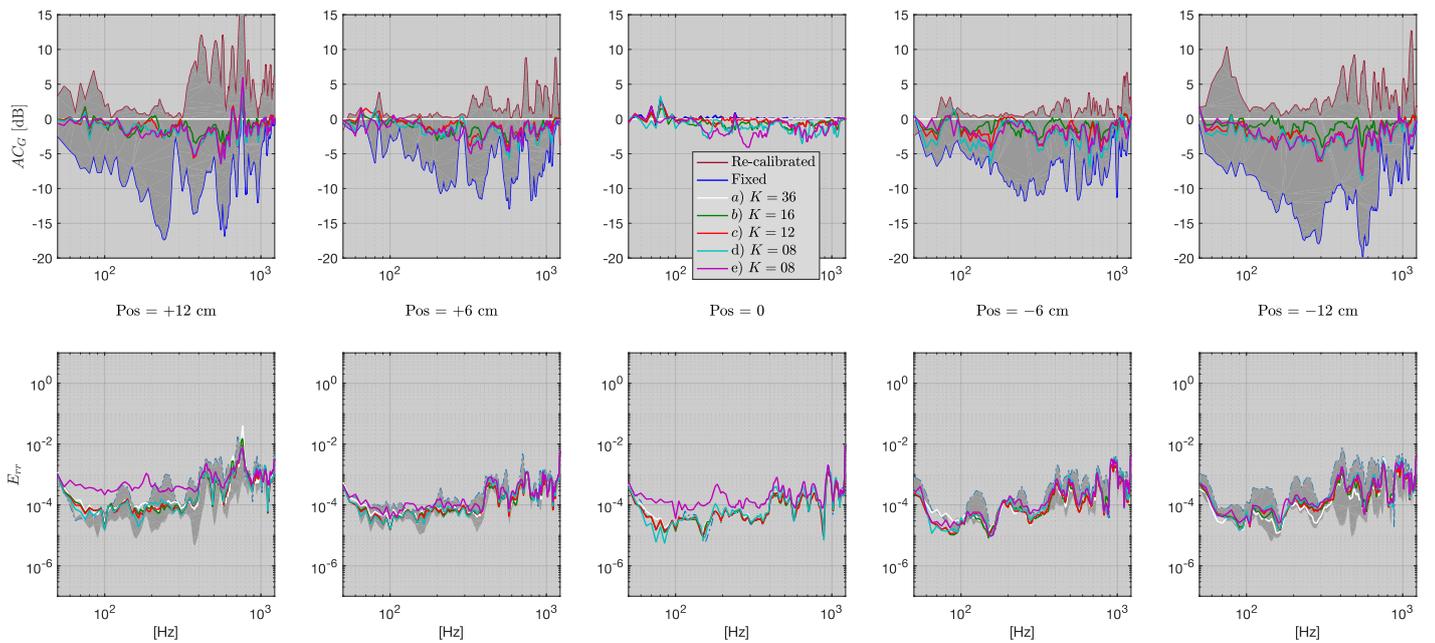


Figure 5.12: FxLMS results with different microphone outlines. Resulting  $AC_G$  and  $E_{rr}$  for symmetric seat positions.

The first conclusion drawn when analysing the resulting  $AC_G$  and  $E_{rr}$  in Fig. 5.12 is that it is possible

to reduce the number of microphones up to a total of  $K = 8$  and the system is still capable of improving the “fixed” configuration regardless of the position of the seat. That is to say, although its performance is not as good as when all the microphones are available -outline *a*) in white- or the Re-calibrated approach, the system is capable of adapting the coefficients with less than  $\frac{1}{4}$  the number of microphones, furthermore, with the exception of a few specific “peaks”, the system is working in the order of 3-4 dB below the  $K = 36$  system.

Regarding the  $AC_G$ , it is possible to observe that for the positions  $\pm 6$  cm all the configurations result in relatively similar performance -close to the 36 Microphone-, although combination *d*) seems to be the worst of all (especially at Pos = +6 cm). When observing the positions  $\pm 12$  cm, outline *e*) is capable of achieving  $AC_G$  of its counterparts with more microphones

Nevertheless, it is necessary to mention that the inaccuracies of the plant are present at the extreme positions, for example at Pos = -12 cm at 550 Hz and 290 Hz, and at Pos +12 at 370 Hz but as the metric is normalised with  $K = 36$  these effects are reduced.

Regarding the  $E_{rr}$ , from outlines *a*), *b*), *c*) and *d*) present very similar results in almost all the positions. However, the only configuration that does not include microphone 7 -outline *e*)- presents a loss in  $E_{rr}$  that can be seen in the position +12 cm but mainly in the centre position. This outline happens to coincide with the only one that does not include the information coming from microphone 7, the closest of all to the loudspeakers.

It is important to analyse in detail what happens at Pos = 0. It was observed in Sec. 5.3.2 that, when using 36 microphones, as the EP is the same as the CP there is no need to adapt the filters so the FxLMS algorithm’s change of the filter coefficients is negligible. However, when using outline *e*) (in Fig. 5.12) both the  $AC_G$  but especially the  $E_{rr}$  are deteriorated in this position. This is due to the fact that, the cost function that minimises the complete zone ( $K = 36$ ) is not the same as the one that minimises only in a few positions ( $K = 8$ ) and the microphone reduction is losing information. In fact, outline *d*) -with microphone 7-, does not present this problem at all, *i.e.* the information missing was the very-near field information coming from microphone 7.

As it was stated in Sec. 4.5.1, the use of virtualisation microphones was inspired by its implementation for ANC algorithms. Some techniques -*e.g.* the remote microphone technique- assume that the primary sound pressure (from the noise source) is the same at both the physical and the virtual microphone [157], which in the near-field does not apply. Nevertheless, there are several other virtualisation algorithms that might be useful for this application [157], their study is listed among the future works of this thesis.

One could argue why not to make the same virtualisation methodology as the one used in the deterministic approaches, but in this cases the TFs between each loudspeaker-microphone pair was available. When using adaptive algorithms, one measures the resulting pressure (to calculate the error) from all the loudspeakers at the same time.

### 5.4.1 Forced FxLMS: ideas from previous chapters

In Secs. 4.5.1 and 5.4, it was observed that the essential microphone to maintain a low  $E_{rr}$  is microphone 7 (see Fig. 5.11), in other words, one of the main contributors to the reproduction part of the problem. Throughout Chapter 3 it was observed that the filters associated with the sources used to define the desired pressures in the  $BZ$  ( $P_R$  and  $P_L$ ) were, regardless of the environment, almost perfect all-pass filters. Even more, in Sec. 4.6 it was showed that it is possible to fix these filters in the calculation with minimum (almost zero) reverberations in the overall performance, guaranteeing a low  $E_{rr}$  even under unrealistic levels of noise [163].

Moreover, in Fig. 5.12 one observed that when only using microphones 4 and 9, the only part of the performance that is deteriorated is the  $E_{rr}$ , the  $AC_G$  is even comparable to configurations with more microphones, so one could assume that microphone 7 are big contributors to the  $BZ$ ’s part of the problem but not so essential for the  $DZ$ .

On the one hand, one could propose to design an asymmetrical microphone outline by only leaving microphone 7 close to the  $BZ$  (in the passenger seat) and using a different microphone outline in the  $DZ$

(driver seat). However, this limits to have the same performance when the zones are inverted. On the other hand, if the reproduction part of the problem is already solved (see Sec.4.6.5), the need of the information provided by microphones 7 might become negligible.

Therefore, one is going to extend the idea of the FPM to the FxLMS algorithm. Moreover, rather than modifying the target pressure as done for the FxLMS in Sec. 4.6, one is going to “Fix” the response of the filters associated with  $P_R$  and  $P_L$  (two  $\delta[n]$ s from  $\mathbf{w}[0]$ ) for all the EP. In other words, one is going to “adapt” six out of the eight filters (to reduce the number of coefficients to adapt by 25%).

It is worth mentioning that by doing this -as it was reviewed for the FPM-, one is not ignoring the fact that the PSZ is a twofold problem because the information of both the  $BZ$  and  $DZ$  is still present [163]. As it was discussed in Sec. 4.6, the fact that one is forcing some filters in the calculation can be interpreted either as feeding a part of the solution to the algorithm or adding the forced filters as a constraint [163].

To illustrate this novel approach, one is going to use the configurations listed in Table 5.6, note that the number of adaptive  $\mathbf{w}$  are listed assuming only one signal.

Id	$K$	Mic. outline	$\mathbf{w}$ adapted	$\mu$
a)	36	1 to 9	8	$5 \times 10^{-5}$
b)	12	4, 7 and 9	8	$8.5 \times 10^{-5}$
c)	8	4 and 9	8	$1 \times 10^{-4}$
d)	8	4 and 9	6	$1 \times 10^{-4}$

Table 5.6: FxLMS microphone outlines with corresponding step-size parameters for the Forced FxLMS.

One proceeded to calculate the  $E_{rr}$  and the  $AC_G$  using as a reference the FxLMS when  $K = 36$  microphones are available. Moreover, outline a) (See table 5.6) is used to calculate the  $AC_G$ .

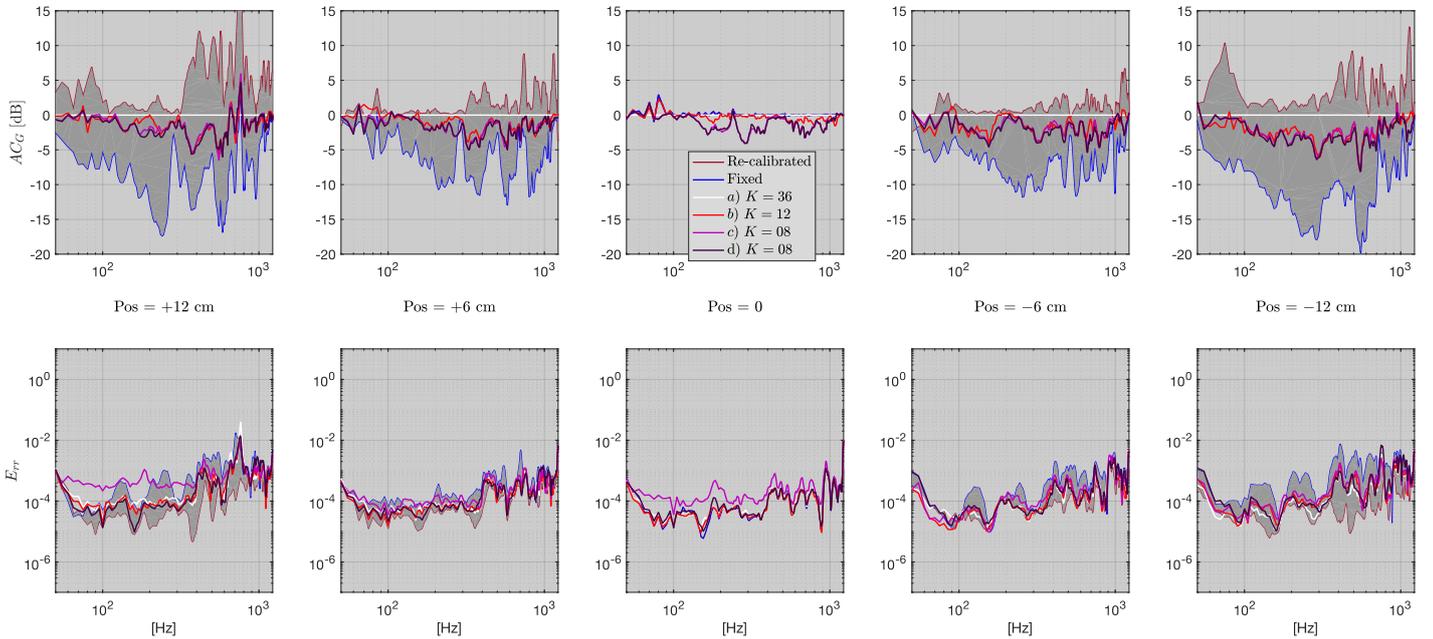


Figure 5.13: Implementation of the forced FxLMS algorithm, compared to the FxLMS with different microphone outlines from Table 5.6. Resulting  $AC_G$  and  $E_{rr}$  for symmetric seat positions.

When analysing Fig. 5.13 it is possible to observe that the hypothesis that microphone 7 was mainly contributing to the reproduction of the problem is validated. By fixing the responses of  $P_R$  and  $P_L$  one “solves” that part of the problem and the algorithm mainly focuses on the attenuation in the  $DZ$ .

In Fig. 5.13, the  $E_{rr}$  is “brought back” to the levels of all the configurations that have microphone 7 available without compromising the  $AC_G$ . Furthermore, by reducing the number of filters to adapt there

is also a reduction of the computational costs.

## 5.5 Different types of acoustic changes

With the conclusions from Sections 5.3 and 5.4 one proceeded to study the implementation of the forced FxLMS for the three studied acoustic modifications (symmetric seat positioning, headrest positioning and addition of passengers in the rear seats). Moreover, all three acoustic modifications assume only one CP at  $\text{Pos} = 0$  and the availability of  $K = 8$  (microphones 4 and 9 in Fig. 5.11).

All the parameters of the adaptive algorithm are listed in the following Table.

Forced FxLMS parameters	
Sampling Frequency	2450 Hz
Filter Coefficients	$M = 200$
Delay	20 samples
CP	Car cabin: seats at $\text{Pos} = 0$ cm
EP	Three different types of changes.
$\psi_{dBZ}$ filter (BZ)	From each EP ( $P_L$ and $P_R$ driven in phase)
$\psi_{dDZ}$ filter (DZ)	Vector of zeros ( $\mathbf{0}$ )
Input	10 s random signal
$\mathbf{w}[0]$	$\mathbf{w}_{CP}$
Plant approx.	IRs from CP
Step-size param.	$\mu = 1 \times 10^{-4}$
Mics. available	$K = 8$ (microphones 4 and 9 in Fig. 5.11)
Mics. evaluation	$K = 36$

Table 5.7: Summary of the forced FxLMS parameters for the study of different types of acoustic changes.

In order to compare with previous methods, one proceeded to plot the resulting  $AC$  and  $E_{rr}$  the results of the adaptive algorithm as a function of the re-calibrated (using the BACCPM) and the fixed. Firstly, on is going to evaluate the resulting metrics when both seats are moved symmetrically (see Fig. 5.2).

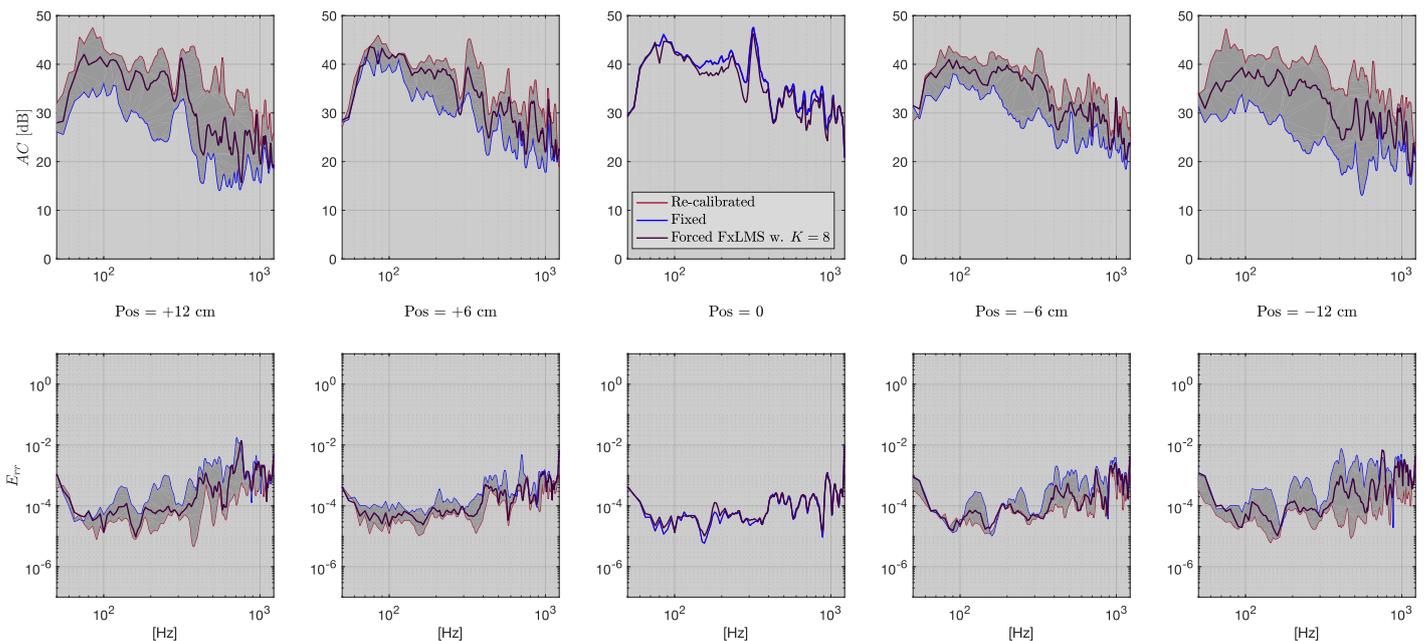


Figure 5.14: Resulting  $AC$  and  $E_{rr}$  for symmetric seat positions, using the Forced FxLMS with  $K = 8$ .

From the results in Fig. 5.14 one can state that the system is capable of improving the performance of the system in all five seat positions. Moreover, the algorithm is capable of improving the  $AC$ , maintaining a low  $E_{rr}$  in the  $BZ$ , by only knowing an approximation of the plant. However, as it was observed in Sec. 5.3.2, the limitations of the plant's accuracy compromise the performance of the system at  $Pos = \pm 12$  cm. Furthermore, there is an important drop in  $AC$  at  $+12$  cm after 400 Hz where the system cannot fulfil the separation requirements. Therefore, the original limitation of having only one CP for the complete cabin might be insufficient (after 10 seconds) when strong acoustic modifications are present.

Regarding the  $E_{rr}$ , the FxLMS does not present any major complications, in most of the cases the system is capable of maintaining a low deterioration in the  $BZ$ . There is a deterioration at  $Pos = +12$  above 600 Hz where the algorithm achieves almost the same level as the fixed approach. This deterioration might be due to the inaccuracy of the plant model at this position, especially for higher frequencies.

As in the deterministic approaches in Sec. 3.4, the higher the frequency the more unstable the system becomes even at  $Pos = \pm 6$  cm. It is possible to observe in Fig. 5.14 that the higher the frequency, in spite of implementing the near field sources, is harder for the system to improve the  $AC$ . This can be seen that for some frequencies the FxLMS falls to the level of the fixed approach (*e.g.* around 700 Hz or above 900Hz). Nevertheless, the system is still capable of maintaining the general performance.

For completeness, in Appendix E other seat-movements are analysed employing also the forced FxLMS algorithm with the same definitions.

On the other hand, one proceeded to study the resulting  $AC$  and  $E_{rr}$  when there is a modification in the headrest height of  $+4$  cm (as in Sec.4.3.2).

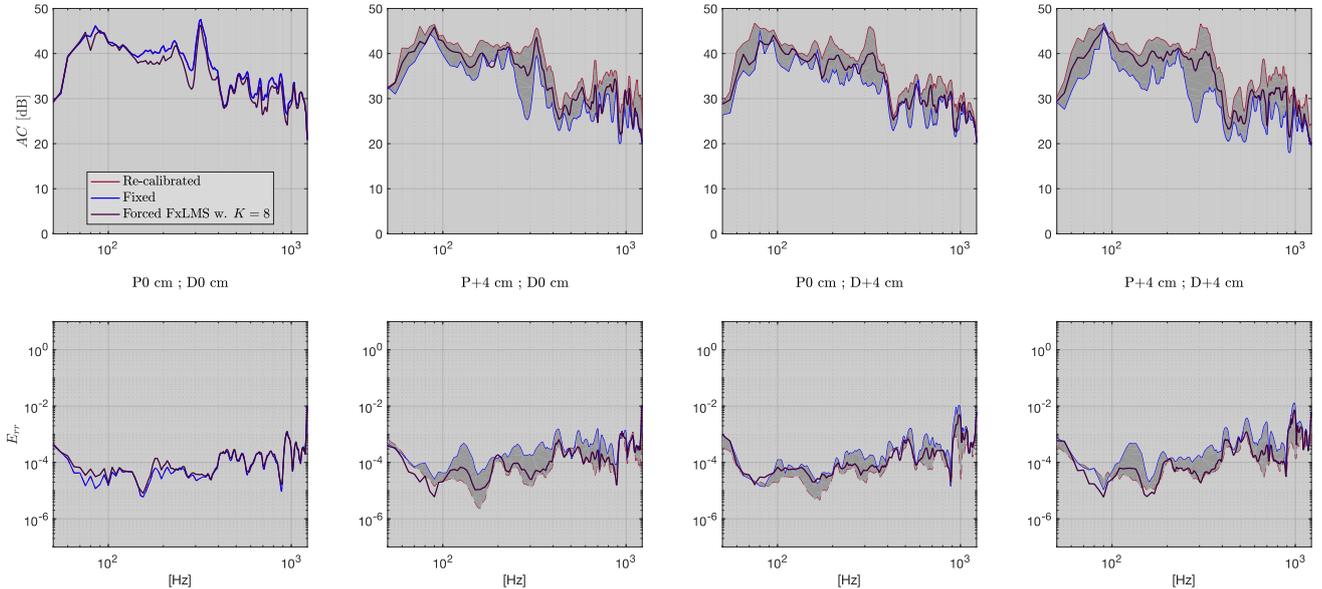


Figure 5.15: Resulting  $AC$  and  $E_{rr}$  for different headrest positions, using the Forced FxLMS with  $K = 8$ .

As in the symmetric seat-position, from Fig. 5.15 one can also state that the system is capable of improving the performance of the system regardless the position of the headrest. Unlike the extreme seat-positions, the plant model when modifying the headrest positioning seems to be accurate enough to allow the algorithm to maintain a performance of 25 dB for the complete bandwidth of interest (except for a drop at  $\simeq 950$  Hz).

Lastly, one proceeded to analyse the resulting  $AC$  and  $E_{rr}$  when passengers are present in the rear seats (as in Sec.4.3.3).

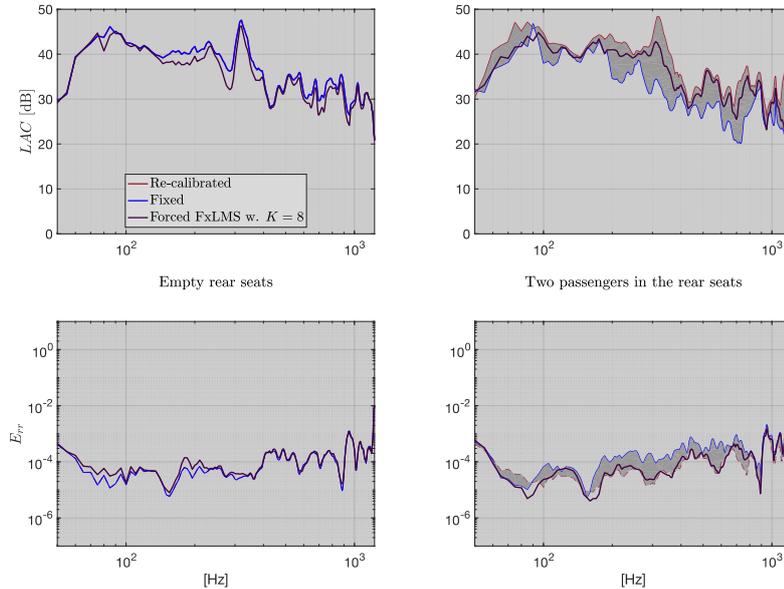


Figure 5.16: Resulting  $AC$  and  $E_{rr}$  when adding passengers in the rear seats, using the Forced FxLMS with  $K = 8$ .

As in the two previous acoustic modifications, in Fig. 5.16 the forced FxLMS is capable of improving the performance of the system when passengers are present in the rear-seats. Under this type of acoustic modification, the adaptive algorithm is capable of achieving almost the re-calibrated performance.

## 5.6 Discussion Adaptive PSZ

The aim of this chapter was to take the first steps towards the use of iterative algorithms to adapt the PSZ coefficients to unknown acoustic modifications that might happen in a car cabin. The core of its application is to take advantage of the fact that these self-designing algorithms can re-calculate their filters by only knowing an approximation/model of the real plant.

Among the many different adaptive algorithms available, one chose the FxLMS not only because it is considered as the benchmark to compare adaptive filter algorithms, but because it requires considerably less computational complexity than its counterparts (*e.g.* FxRLS).

In Sec. 5.3.2, one of the known-problems of the FxLMS algorithm is its frequency-dependent convergence behaviour -something thoroughly studied in the literature (see Sec. 5.1). To overcome this complication without increasing the algorithm's complexity, one proposed to define as initial coefficients the filters calculated using a deterministic approach. In other words, rather than using the FxLMS algorithm to calculate the filters "from scratch", to use the adaptive algorithm to modify filter coefficients that are optimal under different acoustic conditions (at the CP).

On the other hand, although the FxLMS improves the performance of the system at every seat-position, it was observed that the plant model also presents some limitations. It was observed in Fig. 5.8 that the algorithm approaches the optimal filter combination when the seats are at  $\pm 6$  cm, but in the extreme positions  $\pm 12$  cm, where the plant is more different it is not as efficient.

On the other hand, in Sec. 5.3.3 it was analysed that to achieve the smallest values of  $E_{rr}$  it is necessary to have a definition of the desired filter vector ( $\psi_d$ ) at each EP otherwise the algorithm attempts to reproduce the desired from CP -which differs from EP. This will require a  $\psi_d$  identification stage where the filter's IRs can be acquired. Nevertheless, it is necessary to mention this can be done in both seats simultaneously due to the fact that  $\psi_d$  corresponds to the sum of the contribution of  $P_R$  and  $P_L$  to the

microphones. Nevertheless, it was highlighted that it is necessary to assist this aspect from a perceptive point of view since the  $E_{rr}$  might be a too harsh indicator.

One also proceeded to analyse the possibility to reduce the number of installed microphones in the car headrest with a successful reduction of the number of installed microphones without compromising the overall performance considerably.

The findings of the proposed FPM were extended to the FxLMS with a noticeable improvement of the  $E_{rr}$  without compromising the  $AC_G$  and reducing the number of coefficients to adapt by 25%.

Lastly, one added to the symmetric seat-position study, the headrest elevation and passengers in the rear seats. The forced FxLMS was implemented using only eight microphones to maintain the system performance in the presence of the aforementioned changes. It was observed that the FxLMS algorithm is capable of improving the performance in all the acoustic conditions, especially when the acoustic modifications were smaller, *i.e.* seat positions  $\pm 6$  cm, headrest modifications and passengers.

Following, a summary of the advantages and disadvantages of the use of the FxLMS to adapt PSZ:

#### Advantages:

- **Small storage:** It is only necessary to store one plant model and one set of filters.
- **Short-calibration:** The system works with only one CP.
- **No anticipation:** There is no need to know in advance the possible EP.
- **No re-acquisition:** No need to stop and re-measure the IRs.

#### Disadvantages:

- **Installed microphones:** It is necessary to install microphones in the cabin, with all their associated material (converter, cables, signal processing, etc.)
- **DSP:** It needs a dedicated signal processing unit in order to calculate the coefficient updates.
- **Feedback Loop:** Instability, speed of convergence, etc.

### Future Work

The aim of this chapter was to take the first step towards the use of adaptive algorithms to adjust the PSZ filters without the need of an accurate knowledge of the acoustic conditions. However, this study must be seen as a proof-of-concept, now it is necessary to implement the results in an on-line manner to fully assess its feasibility. There are some important aspects that need to be considered, such as the computational limitations available for the audio system, the sensitivity to unknown disturbances, considerations about causality, etc.

Furthermore, one of the necessary following steps is to consider real-signals as inputs of the adaptive algorithm (music, voice, etc.) and study their effect on the convergence of the algorithm. In addition, it is important to study the two zone situation. In other words, to generate two  $BZ$ s and two  $DZ$ s simultaneously and analyse if the system can still maintain its performance.

Similarly to what was mentioned in Sec. 4.5 regarding the microphone removal, although encouraging results were obtained, this is still not the real application, a thorough study needs to be performed in order to find the optimal positioning according to the real available locations in the headrest and validate the system using the microphones incorporated in the seat.

In spite of the fact that the FxLMS algorithm was used due to its simplicity, its low computational cost and the fact that it is a test benchmark, there are several improvements that can be done. For example, in this Chapter the step-size parameter was, not only chosen empirically but also was fixed during the complete learning time of the algorithm. However, many authors propose to use variable step-size parameter which can result in faster-convergence, better performance, smaller sensitivity to noise and smaller sensitivity with respect to the system parameters [170, 171]. Therefore, it is necessary to evaluate algorithms with automatic and variable step-size parameter in order to see if it is possible to further improve the system.

Additionally, one of the strongest requirements imposed in this thesis is the fact that only one calibration

position can be acquired at the end of a production line and it was possible to observe the limitation of the plant model -especially in the extreme positions  $\pm 12$  cm. Nowadays, with electrically controlled seats and cameras, it is possible to know their position at all times. Therefore, it might be possible to overcome these limitations using a position-dependent plant model obtained from maybe two or three calibration positions. Furthermore, it would be interesting to perform a study on the “needed-accuracy” of the plant model in define the needed number of model-plants in order to be able to achieve better results at extreme positions.

Although the ideas originally introduced for the FPM in Sec. 4.6 were successfully applied to the FxLMS, one could take this study further by implementing the independent radiation mode control proposed by Gauthier *et al.* [83] to study the relationship between each radiation mode and the reproduction/cancellation filters.

Moreover, it is important to take a moment to compare the use of initial coefficients as departure of the system and the use of a departure penalty as proposed in the aforementioned paper [83], because there are subtle differences. Firstly, both the PM and the FxLMS are methods based only on an error minimisation cost functions which is not the case of the WFS where no-cost function is involved. Secondly, the “WFS-coefficients” (referred by the authors as WFS operators) are introduced as a penalty in the cost function, which resulted in a re-injection of the solution on each iteration. Whereas in this dissertation, the initial coefficients are defined as “departure-conditions”. Nevertheless, it would be necessary to test if the re-introduction of the initial coefficients as a departure penalty will either help or compromise the system’s convergence.

In addition, inspired by Lee *et al.* [132], it could be interesting to take into consideration the time-varying characteristics of the input signal and the human auditory system applied to a similar iterative algorithm by tuning the definition of  $\psi_d$  in order to reduce the perceived level in the *DZ*.

Furthermore, although the use of virtual microphones was ignored for this study, there are many different algorithms that should be considered to, either enlarge the number of microphones sampling the zones or to be able to obtain the pressure field at further microphones [159, 160, 161, 157]. For example, the importance of the contribution of microphone 9 to the problem was already analysed in Sec. 4.5.1, however it might not be possible to achieve this location physically, therefore the use of a virtual microphone at this location might be the solution.

Lastly, the type of acoustic modification studied in this thesis correspond to stationary environments, after a modification in the acoustic environment, it does not change. In such environments, as mentioned by Haykin [124] *the error-performance surface is fixed, and the essential requirement is to process the training data in a step-by-step fashion, seeking the minimum point of the surface and thereby assuring optimum or near-optimum performance.* However, the adaptive algorithms can be also used to track the continually time-varying position of non-stationary environments. Therefore, it is also necessary to study the “tracking-capabilities” of the algorithm to modifications such as loudspeaker wear.

## The balance of the chapter

The main contributions of the chapter are listed below.

- 1) Implementation of the Filtered-x least-mean-squares algorithms to adapt the PSZ filters to “unknown” acoustic modifications in a car cabin.
- 2) Overcome the FxLMS frequency dependent convergence by providing as initial coefficients the filters from a deterministic approach in different acoustic conditions.
  - Patent pending: N°FR2014082.
  - Paper under review in the Journal of the acoustical society of America.
- 3) Study on the microphone reduction.
- 4) Implementation of the forced pressure matching principles into adaptive algorithms.



# Chapter 6

## Conclusions and future work

### 6.1 Conclusions

The aim of this thesis was to design a sound system capable of generating personal sound zones in a car cabin. Throughout the previous chapters, one attempted to illustrate the research path undertaken from the simplest simulations all the way to a prototype in an acoustically variable car cabin.

From the beginning of the dissertation, one highlighted the requirements and constraints agreed with the manufacturer. Nevertheless, with the contribution of previous works and experimental results, these requirements were questioned and polished.

Different methods were contextualised according to the aforementioned requirements and it was decided that the most suitable candidates were the constrained optimisation methods. This dissertation was divided in two main parts: the invariant acoustic part (Chapters 2 and 3) and the variable acoustic part (Chapters 4 and 5).

#### 6.1.1 Invariant acoustics

In this part, the constrained optimisation methods were implemented using a deterministic approach. It was reviewed that the performance of the overall system depends on the definition of the cost function as well as the acoustical problem to solve (definition of the zones, field constraint, location of the sources, acoustic environment, etc.).

The zones to control were defined within the original requirements; the dark zone (*DZ*) around the driver's ears and the bright zone (*BZ*) around the front passenger's ears. As cost function, it was decided to employ a particular case of the pressure matching methods known as the combined solution. For this method, the minimisation of the cost function allows to reduce the energy sent to the *DZ* as much as possible whilst reproducing a desired pressure in the *BZ*.

As reference system and starting point of the research, one chose to employ two-loudspeaker headrests on each seat ( $L = 4$ ). Therefore, one chose to define the desired pressure in the *BZ* as the natural response of the headrest in that seat. Moreover, it was observed that the "classical" headrest-loudspeakers system was incapable of fulfilling the wanted acoustic contrast (*AC*). Therefore, the focus was put on increasing the number of loudspeakers in order to achieve a larger *AC* deteriorating as little as possible the reproduction in the *BZ*.

A series of simulations and experiments were carried out in order to design and test the new personal sound zones system. These were organised according to their "level of acoustic complexity", from theoretical monopoles in free-field to a prototype in a car cabin.

The final system was composed of the classical headrest system (two sources per seat), to which two inner sources (between the seats) and two outer sources were added, giving a total of  $L = 8$ . In spite of the fact that the contribution of the outer sources proved to be not so important in the anechoic chamber, under more complex acoustic environments (especially in the car cabin), it plays a major role in increasing

the  $AC$  as well as reducing the reproduction deterioration. The final system achieved a mean  $AC$  of 34.7 dB with a mean  $E_{rr}$  of  $1.69 \times 10^{-4}$  over the studied bandwidth in a car cabin.

Among the main concerns of the systems based on CO are: the condition number (to have an ill-conditioned problem) and the effort (demand to the array something that it cannot do). The approach of employing near-field sources, the custom location of the added sources, as well as the definition of the desired pressure (the natural response of the system) resulted in an “acoustically well-posed” problem. In other words, in all the experiments performed in this thesis the effort ( $E_{ff}$ ) was kept below 5 dB and the condition number did not present any complications.

Other complications one can encounter when implementing CO are the resulting filter’s time-shape (their causality, presence of pre-ringing, etc.). Due to the distribution of loudspeakers (near-field of the  $BZ$ ) and the definition of the desired pressure, the filters contributing the most to the reproduction are almost perfect  $\delta[n]$  functions. This permits to assume that the deterioration in reproduction due to the PSZ processing will be negligible.

On the other hand, two different formulations were considered that fall in the category of the “combined-solution”: in the frequency domain (PM) and in the time domain (BACCPM). Although both formulations resulted in almost identical performances, the resulting filters of the BACCPM presented smaller pre-ringing than the ones obtained with the PM. However, the time domain was more than 1000 times slower to obtain the resulting filters than the frequency domain.

One of the key problematic of employing deterministic methods is that they assume that the acoustic environment is time invariant, *i.e.* that the IRs measured are always valid. However, a car cabin is an acoustically-variable environment, *i.e.* many modifications can happen in a car cabin that modify the acoustic response of the system.

## 6.1.2 Variable acoustics

Therefore, one proceeded to evaluate experimentally the sensitivity of the system to “typical-modifications” that can happen in a car cabin. Using the prototype tested in all the previous experiments, three different modifications were chosen as the most likely to occur: modifications in the position of the seats, elevation of the headrest and the incorporation of passengers in the back seats.

The acoustic environment has an important role on the system’s performance. Therefore, the system was tested taking each-modification as a “new-acoustic” environment, *i.e.* going through all the deterministic steps again (referred in this dissertation as re-calibrated approach). In this situation, the system designed in previous chapters was capable of fulfilling the original constraints regardless of the acoustic modification introduced (regardless of the seat positioning, the headrest positioning and the presence of passengers in the rear seats) if a dedicated set of filters was calculated for each “new-acoustic environment”. Conversely, when only one filter combination was applied in all the conditions (referred to as the Fixed approach), the loss of  $AC$  reached  $\geq 15$  dB and, in some cases, even worse performance than not applying any processing at all.

Therefore, the challenge was to find the way to obtain the “optimal” filter combination when there is a modification in the car cabin. A summary of the different strategies proposed in this thesis is presented in the following block diagram, with their advantages and disadvantages.

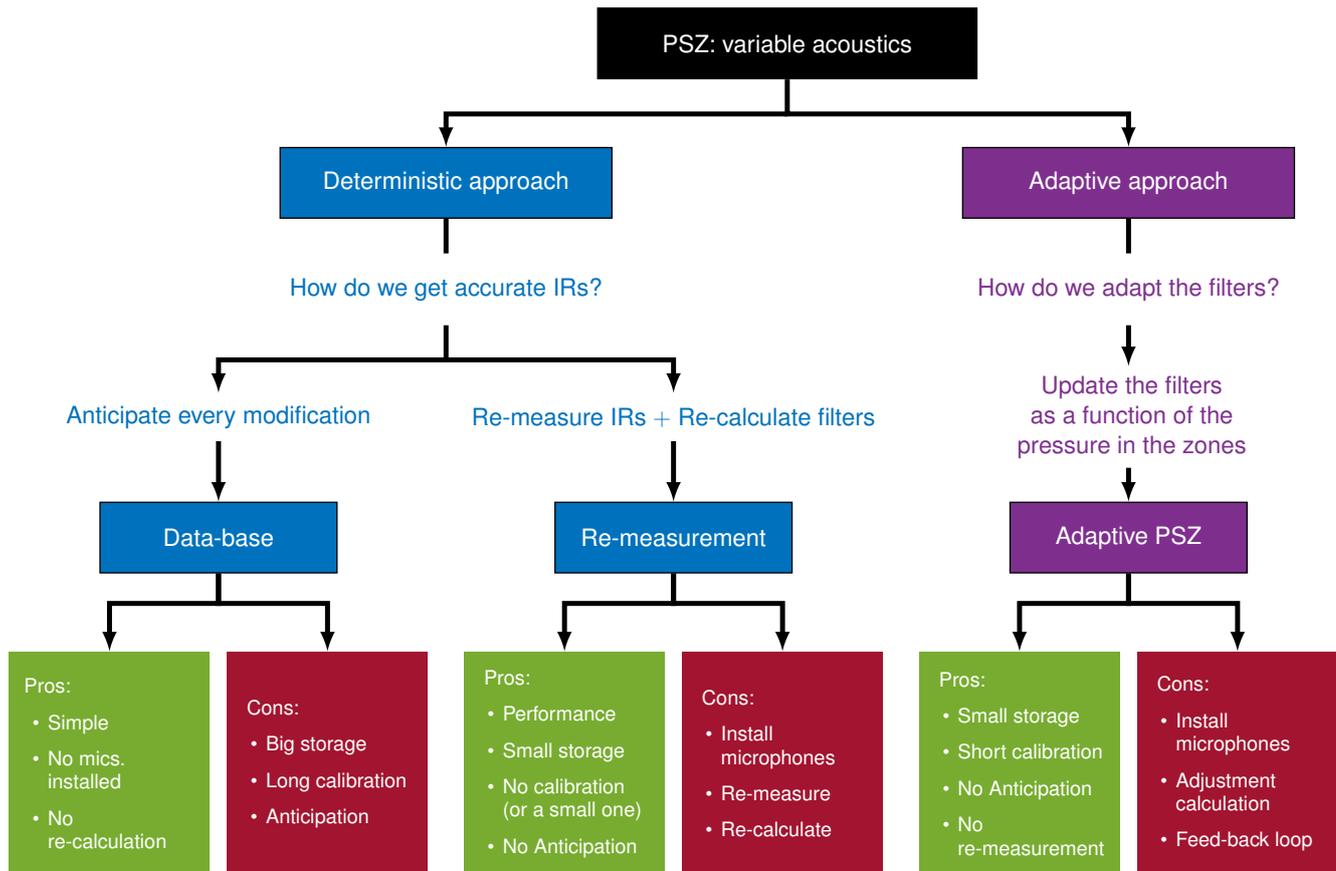


Figure 6.1: Block diagram of strategies used in this dissertation to overcome the acoustic variability in a car cabin.

The methods proposed in this thesis are separated in two: the ones derived from the deterministic approach and one implementing an adaptive approach (see Fig. 6.1).

### Deterministic approach

Two solutions directly derived from the deterministic approaches were considered: the data-base solution and the re-measurement solution. The challenge of these two solutions is how to obtain an accurate characterisation of the acoustic condition, to calculate the optimal filters, to achieve the best performance the system can provide.

The first option considered was the one referred to as data-base solution. This solution consists in anticipating every possible acoustic modification that can happen in a car cabin in the production line and re-run the deterministic steps for each modification. All the optimal filters are stored and chosen according to other sensors (camera, electrical position of the seat, etc.). Assuming that one could anticipate every change in a car cabin, it is necessary to have a dedicated storage unit installed in a car capable of storing all the filter-combinations. Additionally, in the production line, it is necessary to run an extensive calibration procedure going through the deterministic steps.

The second option considered was the re-measurement solution. This solution is based on the premise that, when a loss of performance is detected, one stops the entertainment system and re-runs the deterministic approach steps “on the go”. One of the main challenges of this solution is that, in order to measure the IRs, it is necessary to install microphones in the car cabin. As it is not possible in practice to fully characterise the zones ( $K = 36$ , located close to the heads), a study was performed in order to reduce the number of microphones. Results showed that it is possible to reduce the number of microphones to the point of using only two per ear ( $K = 8$ ) almost without any performance deterioration.

In order to make these two solutions more viable for the car application, the forced pressure matching method was introduced. This method is a rethinking of the pressure matching method when the responses of some filters are imposed in the calculation. The use of the FPM for the application in hand, reduces the amount of storage needed by the database-solution by 25% without compromising the performance. Additionally, the FPM reduces the calculation time of the filters by  $\simeq 21\%$  without compromising the performance.

### Adaptive approach

As an alternative to the deterministic approach, the use of adaptive filter algorithms is proposed to overcome the acoustic-variability of the car cabin. The core of these methods is to iteratively adapt the filters as a function of the resulting pressure in the zones. In this dissertation, the FxLMS algorithm was studied due to its simplicity and low-computational cost.

The key feature of the FxLMS implementation is that, in order to calculate the filters, it only requires an imperfect knowledge of the car cabin (plant model). Therefore, it was hypothesised that the IRs from one acoustic conditions (or as referred in this thesis, one calibration position) are enough to calculate the filters under any of the studied modifications.

Furthermore, one of the problems of the FxLMS algorithm is that it presents a frequency dependent convergence, compromising the performance of the system especially in the lower frequencies. Hence, in this dissertation one proposed to overcome this limitation by using as starting point of the adaptive algorithm (as initial coefficients) the optimal filters calculated at the calibration conditions using a deterministic approach (run the deterministic steps only once).

When all  $K = 36$  microphones are available, the FxLMS algorithm is capable of improving the fixed performance in all the studied modifications. The use of a single plant model showed to be capable of achieving levels of  $AC$  and  $E_{rr}$  similar to the re-calibrated approach when relatively small modifications were present. In more extreme positions, the system still improves the performance but it did not achieve the re-calibrated levels.

Nevertheless, as in the re-measurement methods, the solution studied requires to install microphones in the car cabin in order to calculate the filters. Therefore, the system was successfully implemented the FxLMS algorithm only using eight microphones ( $K = 8$ ).

Lastly, it was also showed that it is possible to extend the principles of the FPM to the FxLMS algorithm, not only to reduce the number of coefficients to adapt, but to improve the  $E_{rr}$  for some microphone configurations.

## 6.2 Future work

The next-step of this thesis is the study of a full-band system (20 Hz - 20 kHz) using the findings of this thesis in conjunction with other technologies. It is necessary to evaluate, not only how to combine the different solutions, but also how to optimise the compatibility of the different technologies in terms of the number of independent channels, amplifiers, placing, etc.

In addition, it is also necessary to consider the extension of the problems to a multi-zone problem incorporating the back seats. It would be necessary to consider increasing the number of sources as well as assisting the system by incorporating passive-directive sources as in previous studies.

Several aspects concerning the real implementation should be added to the constraints and requirements of this study. For example, signal processing capabilities in a car cabin application, end-of the line requirements (such as available calibration time), compatibility with other technologies (ANC for example), etc.

Another important step that needs to be addressed, is the -real- implementation of microphones in the car cabin. In this thesis it was showed that the microphone reduction is possible but a dedicated study needs to be performed considering embedded microphones in the cabin in order to maximise the

performance of the re-measurement and/or adaptive solutions. Moreover, this study needs to be done considering the use of virtual-microphones to achieve locations that are not physically possible.

In the author's opinion, the best trade-off would be to implement both the re-measurement as well as the adaptive solutions conjointly. In other words, use the strengths of the adaptive algorithm to adjust to small modifications, whilst in the presence of a major drop in performance, run the re-measurement solution in order to adjust the adaptive algorithm with a new plant-model and new set of initial coefficients.

The findings of this thesis regarding PSZ-adaptation to unknown acoustic conditions should be seen as the first steps of a long research topic. Several improvements can be considered in order to continue this particular problematic.

For example, there is one important aspect that needs to be studied in detail, the calibration conditions (plant model, initial coefficients, etc.). In this thesis the simplest scenario was considered; only one acoustic condition (calibration position) for the complete system. However, the best calibration position might not coincide with the  $Pos = 0$  of the cabin, or it might be necessary to make a combination of calibrations to achieve optimal performance, or it could be possible to run a more complex plant identification in order to adjust it as a function of the seat position, etc. Therefore, a thorough analysis should be carried out in order to optimise the calibration stage considering both the performance as well as the end of the production line conditions.

On the other hand, one of the most important aspects missing in this dissertation is the perceptual evaluation of the designed system. In spite of the fact that previous studies consider the limit of 25-29 dB as the target  $AC$ , it has been discussed that the  $AC$  represents a sort-of average over the zone and that it does not necessarily represent the perceived "separation", hence, a thorough perceptive study should be performed in car conditions in order to evaluate the performance of the system. Moreover, although the movements of the passengers are limited due to the seatbelt, another important aspect that should be studied in detail is the sensitivity of the system to movements of the head, especially when re-measuring or adapting the filters.

Furthermore, there is a big unanswered question concerning the  $E_{rr}$  and its relationship with the "transparency" of the system. Although very low levels of  $E_{rr}$  are achieved (and a perfect  $\delta[n]$  is ensured as the reproduction filters), the  $E_{rr}$  also presents an average over the complete zone. In other words, it might be possible to allow a slightly larger deterioration of the zone in order to achieve a quieter  $DZ$ . Furthermore, it would be important to quantify the reproduction requirements as a function of the spatialisation technologies in the market.

Although some informal tests were performed showing that the separation achieved was successful, the effect of movements were not so significant, it is necessary to formalise a series of perceptive tests targeted to the car application, especially considering the driving situation.



# Appendices



# Appendix A

## Headrests in cars

In this appendix, a list of car-applications involving the use of headrest loudspeakers is given.

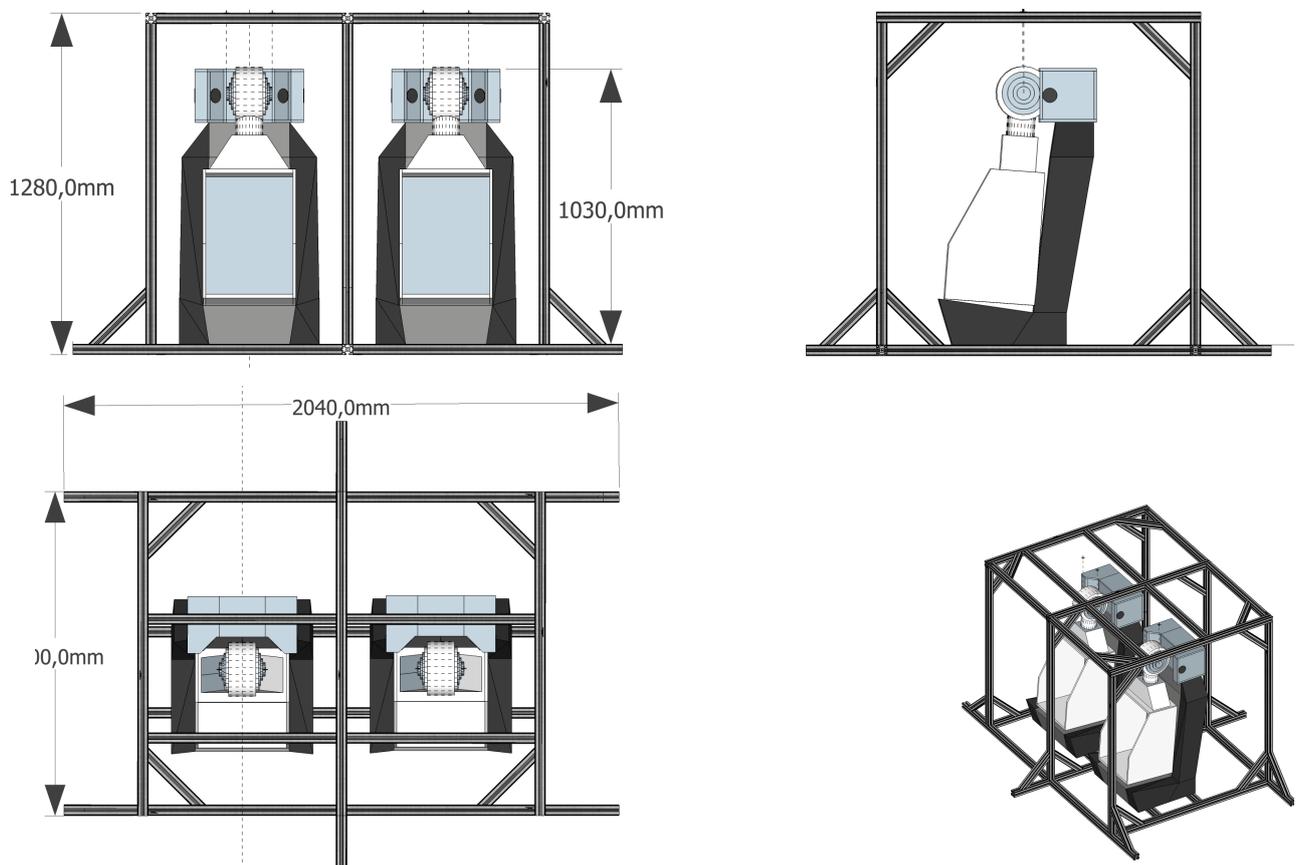
- BOSE:
    - <https://automotive.bose.com/technology-systems/seatcentric-experiences>
    - <https://automotive.bose.com/vehicles/mazda/mx-5-miata>
    - <https://www.fiatcanada.com/en/spider>
  - Bang and Olufsen
    - <https://corporate.bang-olufsen.com/en/partners/automotive/audi/r8>
    - [https://news.harman.com/blog/blog\\_articles-20200629](https://news.harman.com/blog/blog_articles-20200629)
  - AKG
    - <https://www.akg.com/ForCadillac.html>
  - Sennheiser
    - <https://en-us.sennheiser.com/ambeo-mobility>
  - Skoda
    - <https://www.carscoops.com/2018/04/skodas-april-fools-noise-cancelling-headrest-actually-sounds-useful/>
- List of technologies proposed for car applications using headrests:
- Faurecia
    - <https://www.faurecia.com/en/innovation/smart-life-board/immersive-sound-experience>
  - Harman
    - <https://www.individualsoundzones.com>
    - <https://news.harman.com/file?fid=58b96966a138355a0e972fbb>
  - Sony
    - <https://www.sony.net/SonyInfo/vision-s/entertainment.html>

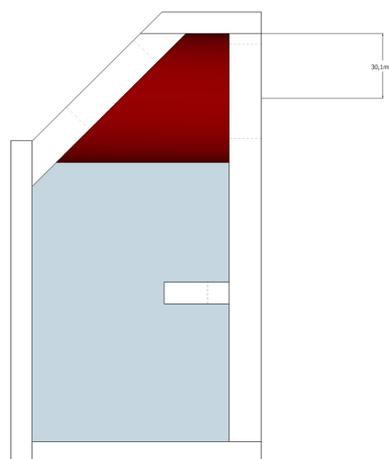
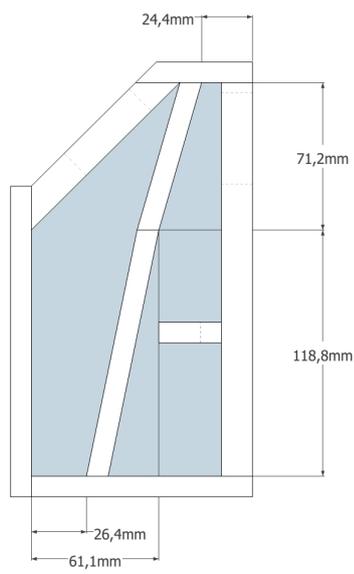
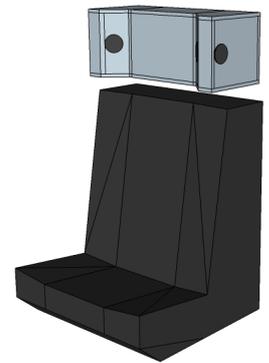
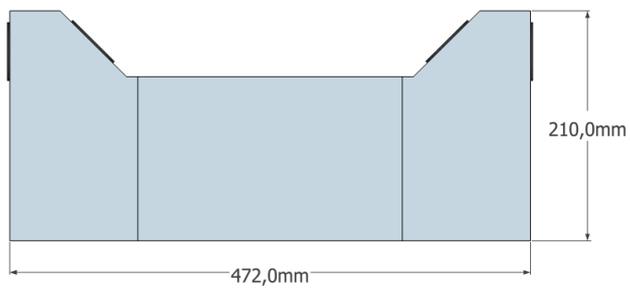
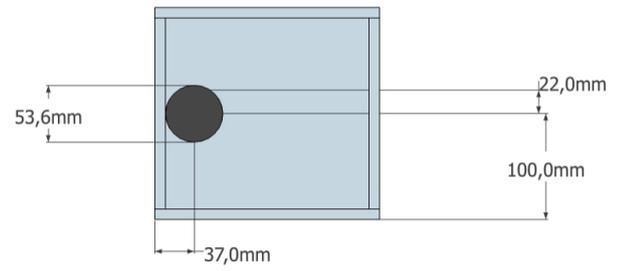
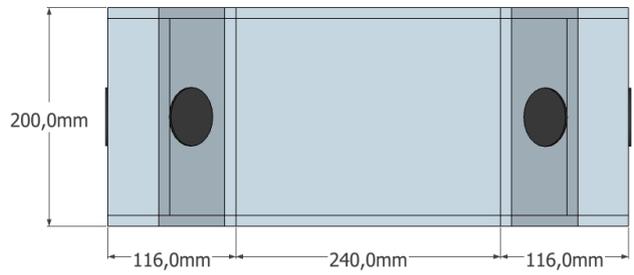


## Appendix B

# Prototype Schematics

In this appendix, the schematics used to build the system studied in this dissertation is introduced.





# Appendix C

## Datasheet

In this appendix, the datasheet of the loudspeaker used for the prototype is presented.



**TEBM35C10-4**  
**Data Sheet**

---

### TEBM35C10-4 Miniature BMR<sup>®</sup> Driver

✓**RoHS**  
COMPLIANT



#### Features

- Full range: 100Hz – 20kHz
- Extremely wide directivity; 180°
- Nominal Impedance: 4 ohm
- Diameter: 52mm (54mm OD max)
- Depth: 25.1mm
- Mass: 51.3g

#### Applications

- Portable speakers
- Sound bars and stands
- Flat TV speakers
- Conference speaker phone

#### Description

The TEBM35C10-4 BMR<sup>®</sup> is an audio drive unit with an extended frequency response and extremely wide directivity. It combines the benefits of Tectonic Elements bending-wave technology and pistonic modes of operation.

The small form-factor is ideally suited for compact products that require a full-range drive unit, room filling sound and a high performance acoustic solution.

**Parameters**

Parameter	Description	min	typ	max	Units
$R_e$	DC resistance	-10%	4.2	+10%	Ohms
$L_e$	Inductance (@ 10kHz)	-	0.05	-	mH
$BL$	Force factor	-	2.03	-	Tm
$f_s$	Resonant frequency	-20%	145	+20%	Hz
$SPL$	Sound Pressure Level @ 1W, 1m	-	80	-	dB
$d_{Drv}$	Voice coil diameter	-	20.4	-	mm
$M_{ms}$	Moving mass	-10%	1.11	+10%	g
$C_{ms}$	Compliance	-10%	1.04	+10%	mmN <sup>-1</sup>
$R_{ms}$	Suspension Loss	-	0.27	-	Nsm <sup>-1</sup>
$X_{mech\ max}$	Maximum coil excursion (p-p)	-	8.0	-	mm
$S_d$	Effective piston area	-	11.04	-	cm <sup>2</sup>
$V_{AS}$	Equivalent volume	-	0.18	-	L
$Q_{ms}$	Mechanical quality factor	-	3.84	-	
$Q_{es}$	Electrical quality factor	-	1.06	-	
$Q_{ts}$	Total quality factor	-	0.83	-	

**Operating conditions**

Condition	Value
Power handling (continuous, weighted pink noise)	10W
Operating temperature range	-20 to 55° C
Audio frequency range	100Hz to 20kHz

**Measured Response – on axis SPL**

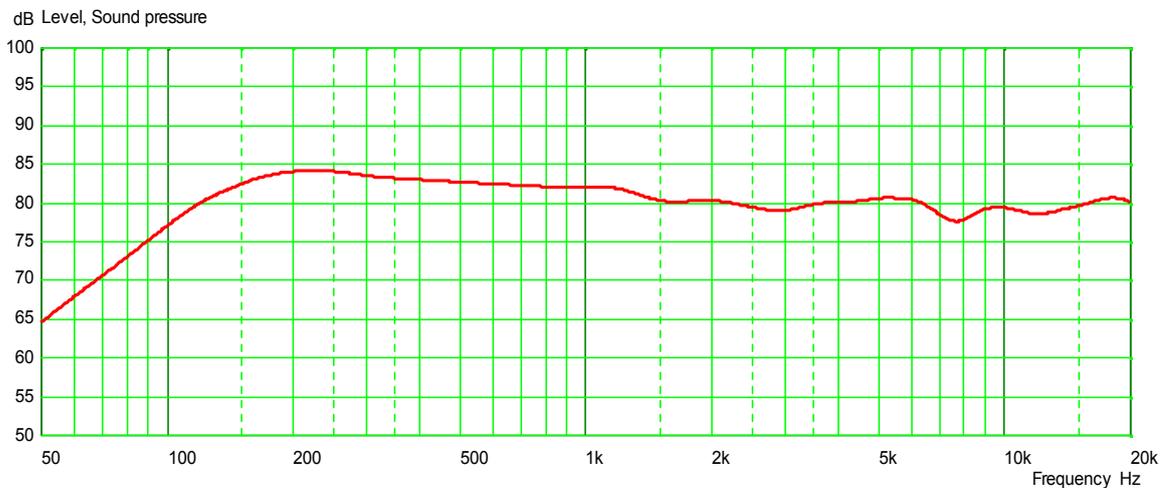


Figure 1: Red: on-axis SPL at 1W/1m (1/3-octave smoothed/spliced/ in room measurement)

### Measured Response – wide angle averaged SPL

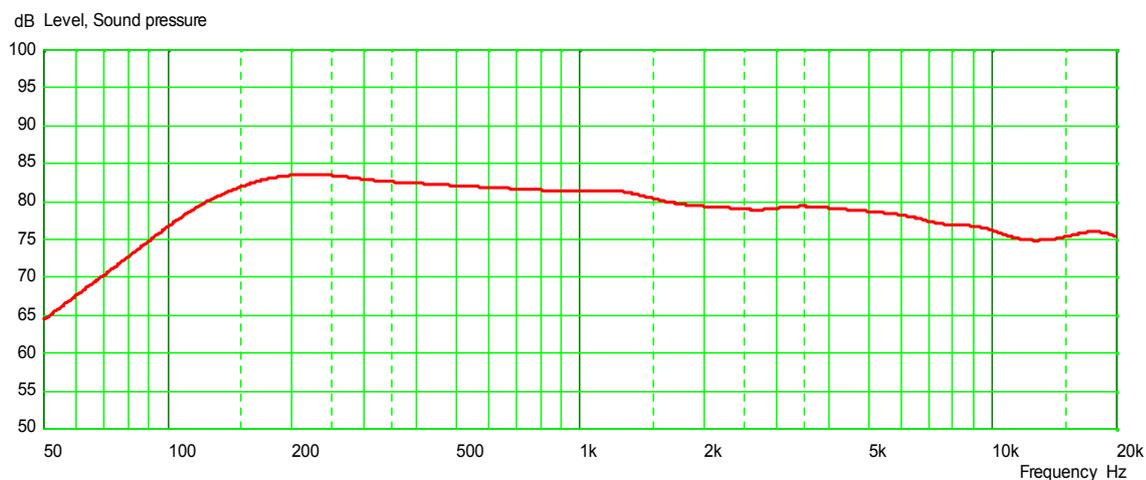


Figure 2: Averaged SPL from measurements across -85 to +85 degrees, 1W/1m, (1/3-octave smoothed/spliced)

### Measured Polar Response

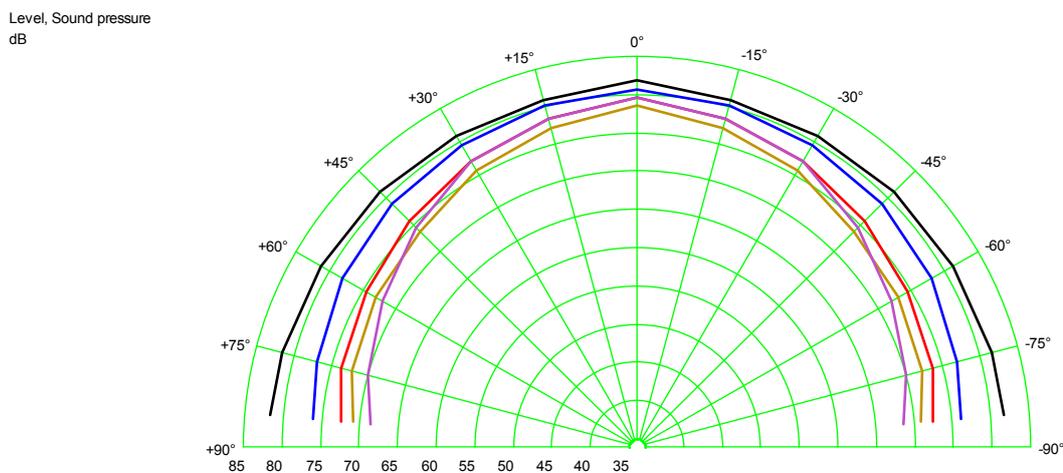
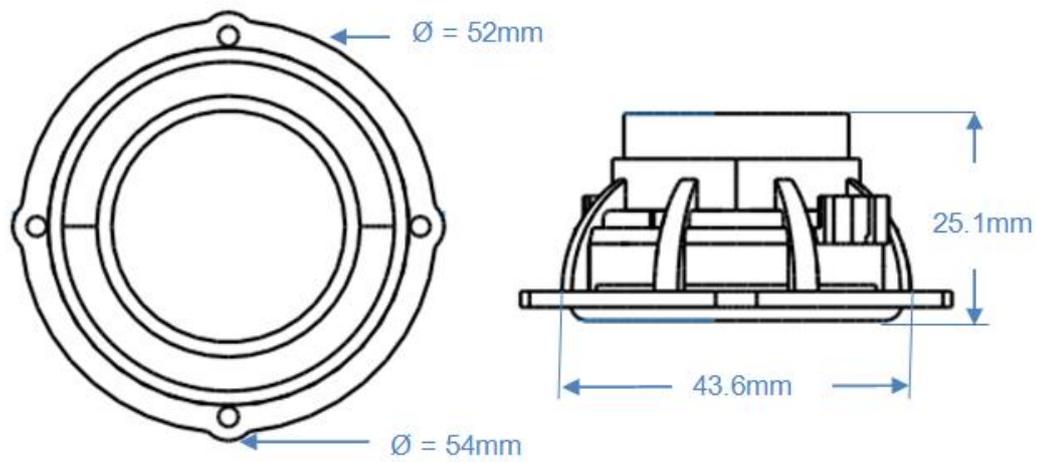


Figure 3: Polar response, angle/dB SPL: Black:1kHz, Blue: 5kHz, Red: 10kHz; Brown: 12kHz; Purple: 15kHz

**Outline Drawing****Figure 4. Nominal dimensions**

*Disclaimer: The information in this Data Sheet is subject to change without notification*

*Please see [www.tectonicelements.com](http://www.tectonicelements.com) for Terms and Conditions of Sale*

[https://www.audiophonics.fr/en/index.php?controller=attachment&id\\_attachment=933](https://www.audiophonics.fr/en/index.php?controller=attachment&id_attachment=933)

## Appendix D

# Online results

In order to validate the use of off-line results, one proceeded to study the use of the algorithms proposed in an on-line manner, *i.e.* driving the system with filtered MLS signals and acquiring the resulting pressures over 48 microphones sampling the zones to control. The system was evaluated in a car cabin as in Sec.3.4.

It is important to mention that both results below are using the frequency and time formulations without any type of regularisation Sec. 2.3.3.4.

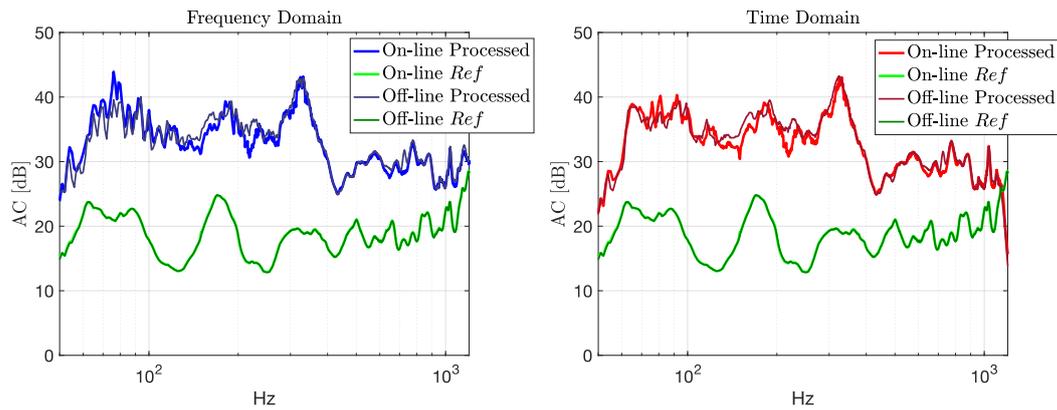


Figure D.1: Online  $AC$  measured in a car cabin using 48 microphones with a 1 Hz resolution

Comparing the on-line and off-line  $AC$  one can conclude that both result in very similar  $AC$ , hence, one decided to continue the work in an off-line manner for simplicity.



## Appendix E

# Crossed Movements Car cabin

In order to complete the analysis showed in the main body of this dissertation, one added to the already-studied symmetric movements (in Chapters 4 and 5) the “crossed movements”. These movements consists in leaving one of the seats in the centre position (Pos=0) and moving the other  $\pm 12$  cm with a 6 cm step (see Fig.E.1).

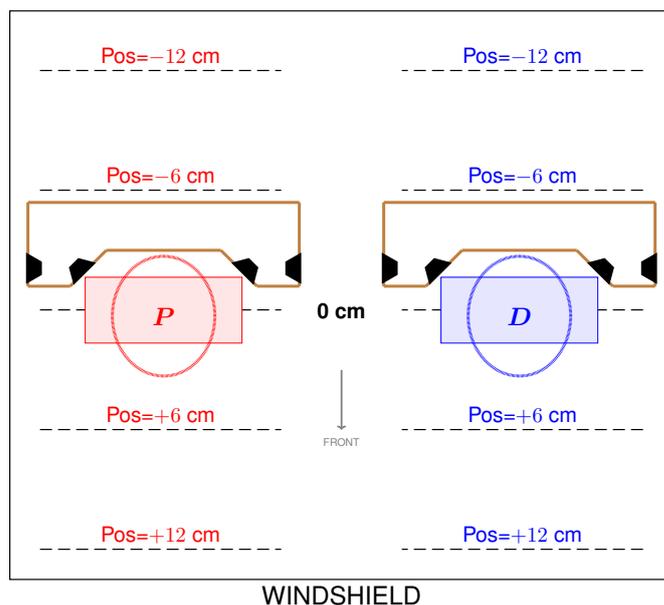


Figure E.1: Schematics: changes in the seat positions (copy of Fig. 4.1a)

Therefore, one proceeded to analyse both the Re-calibrated, the Fixed as well as the implementation of the FxLMS algorithm employing only  $K = 8$  microphones as in Sec. 5.5.

Firstly, one proceeded to analyse the system by only moving the passenger seat, *i.e.* the *BZ*.

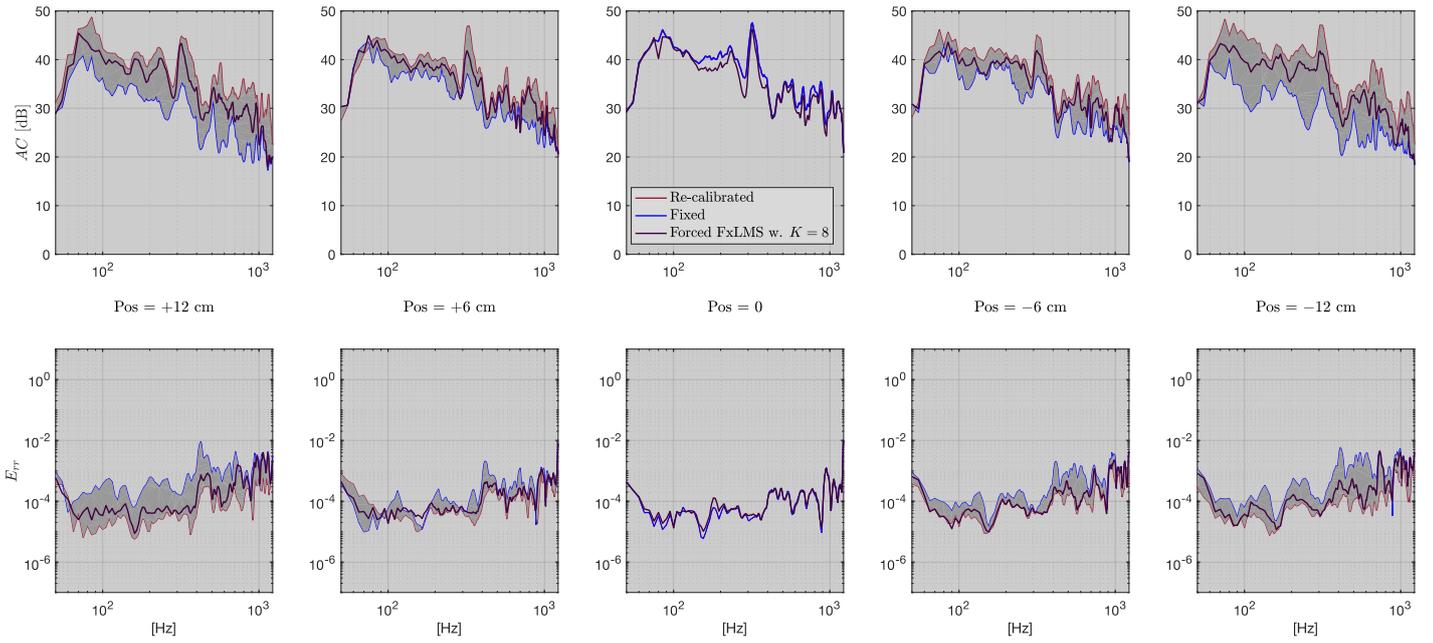


Figure E.2: Resulting  $AC$  and  $E_{rr}$  for crossed changes in the seat position moving only the passenger ( $BZ$ ), using the Forced FxLMS with  $K = 8$ .

Firstly, in Fig.E.2 it is possible to observe that the deterioration of the Fixed approach is smaller than the one obtained when symmetric movements were present (see Fig.5.14). Moreover, the FxLMS algorithm is capable of improving the performance in all the positions and, unlike with the symmetric movements (see Fig.5.14), it is capable of fulfilling the original constraints even when the seat is at  $\pm 12$  cm (at least up to 900 Hz).

On the other hand, one also proceeded to only move the driver seat, *i.e.* the  $DZ$ .

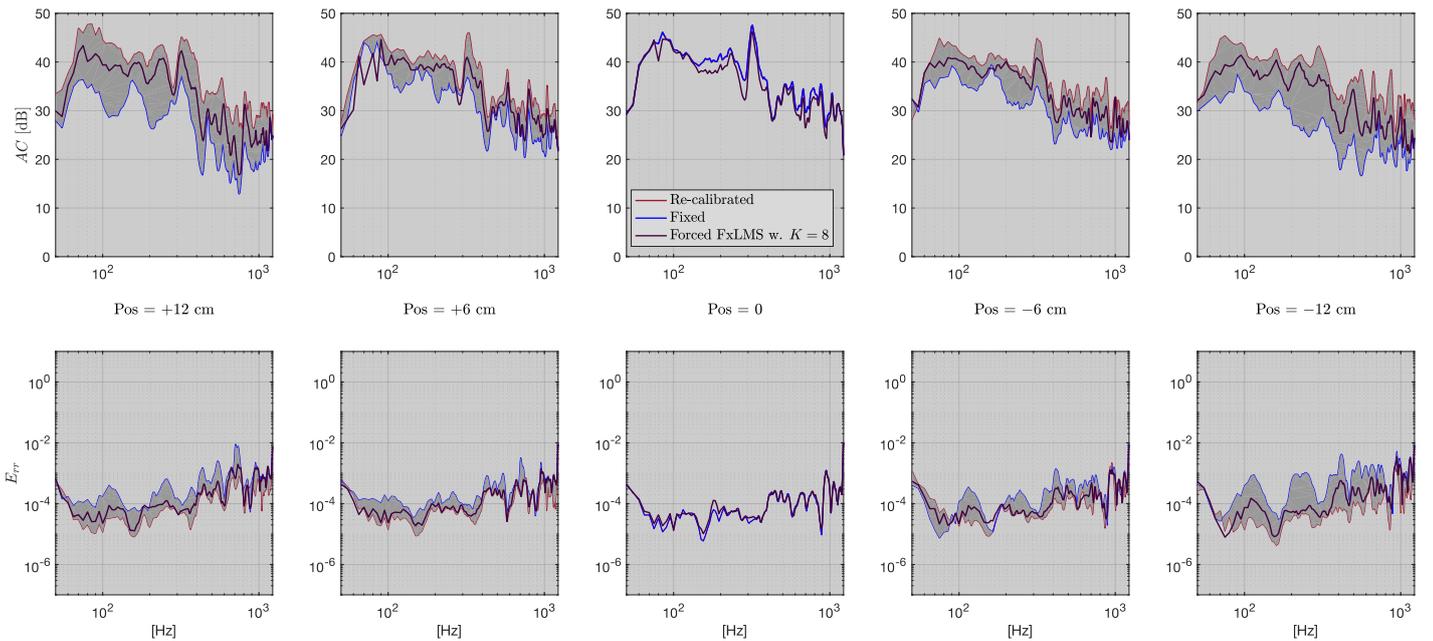


Figure E.3: Resulting  $AC$  and  $E_{rr}$  for crossed changes in the seat position moving only the driver ( $DZ$ ), using the Forced FxLMS with  $K = 8$ .

The results obtained in Fig. E.3 are closer to the ones obtained when symmetric movements were present (see Fig.5.14). The deterioration of the fixed approach is considerably larger at  $\pm 12$  cm and it is harder for the adaptive algorithm to achieve the “optimal” performance (especially at +12 cm).



# Appendix F

## FxLMS: using simulated IRs

In this appendix, the results of the same preliminary study as the one performed in Sec. 5.3.1 is shown but using eight monopole sources in Free-field as in Sec. 3.2.1.

### F.1 Flat responses

Firstly, one proceeded to re-do the preliminary study with theoretical monopole sources in free-field. Four FRF are shown in Fig. F.1 of the same microphone in the four different ears of the passengers when only one loudspeaker ( $P_R$ ) is driven without any processing.

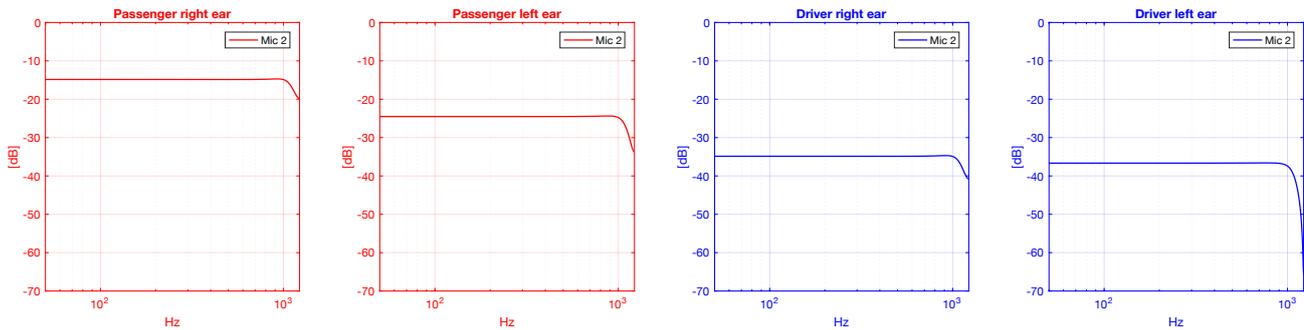


Figure F.1: FRF when only driving loudspeaker  $P_R$  without any processing for the four microphones located at the ears, (microphones 2 in Fig. 4.11).

Using Eq. 5.17 it is possible to calculate the convergence condition, for this evaluation it has been assumed that  $\tilde{\mathbf{R}} = \mathbf{R}$ , which results in  $\mu < 4.1 \times 10^{-3}$ .

In Table F.1 a summary of the definitions used for the FxLMS is presented.

<b>FxLMS parameters: preliminary study</b>	
Sampling Frequency	2450 Hz
Filter Coefficients	$M = 200$
Delay	20 samples (comprised in $\psi_d$ )
CP	Monopoles in free-field
EP	Monopoles in free-field
$\psi_{dBZ}$ filter (BZ)	As in the deterministic approaches = $\mathbf{p}_{dBZ}$ ( $P_L$ and $P_R$ driven in phase )
$\psi_{dDZ}$ filter (DZ)	Vector of zeros ( $\mathbf{0}$ )
Input	10 s random signal
Filters at $n = 0$	$\mathbf{w}[0] = \mathbf{0}$
Plant approx.	$\tilde{\mathbf{R}} = \mathbf{R}$
Step-size param.	$\mu = [3.5 \times 10^{-3} - 5 \times 10^{-4}]$
Mics. available	$K = 36$

Table F.1: Summary of the FxLMS parameters for the study of the  $\mu$  effects using simulated IRs.

Fig. F.2 shows the learning curve of the algorithm using simulated IRs for different values of  $\mu$ .

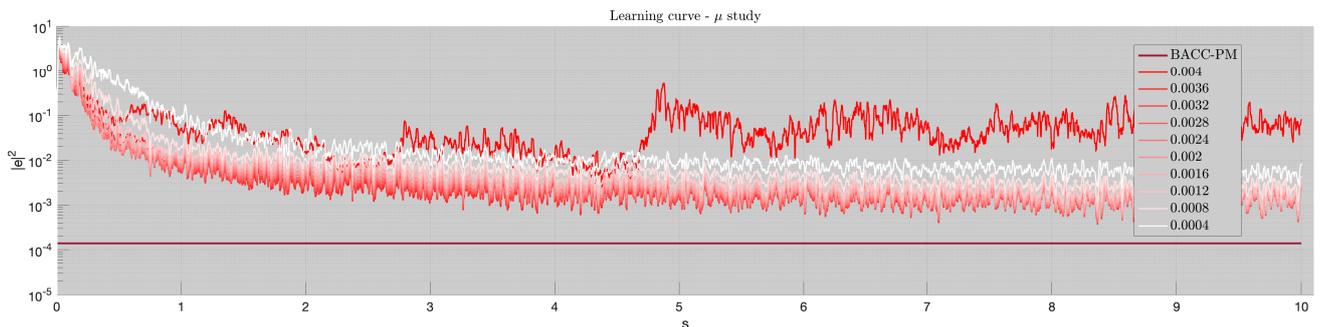


Figure F.2: Learning curve for different  $\mu$  values at Pos = 0

In order to extend these conclusions to the PSZ literature, one proceeded to plot both the resulting  $AC$  and  $E_{rr}$  using the last filter combination obtained after 10 s for different values of  $\mu$ .

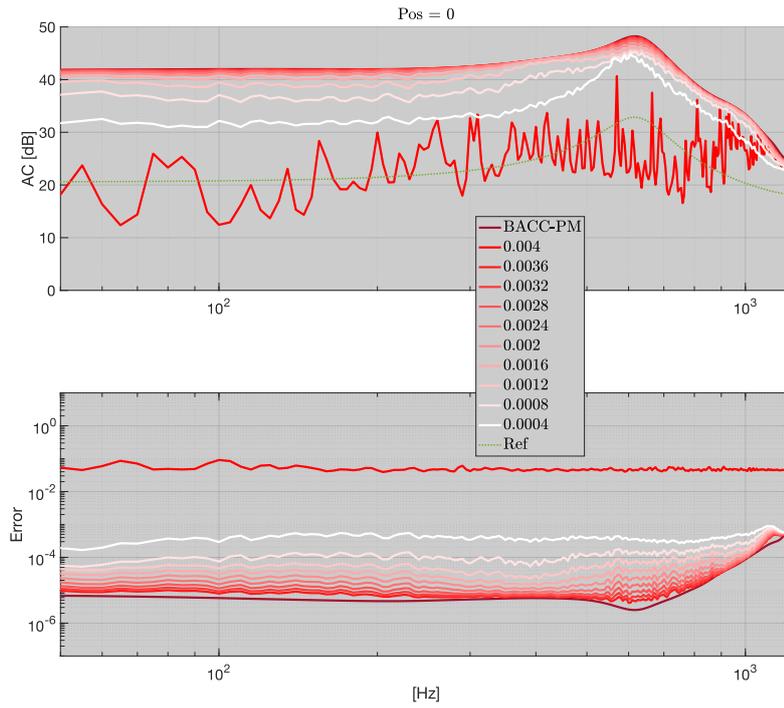


Figure F.3: Results: Metrics at Pos = 0 using FxLMS after 10 s

It is possible to observe in Fig.F.3 that, for the larger values of  $\mu$ , as long as the system is converges, the frequency dependent characteristics of the FxLMS observed in Sec. 5.3.1 are not present

## F.2 High-pass responses

In order to evaluate the frequency-dependent convergence of the FxLMS a second-order high-pass filter is applied to the theoretical monopoles used in the previous section.

To simulate the loudspeaker’s FRFs, a second-order high-pass filter is used with a cut-off frequency of 150 Hz and a  $Q$ ts of 0.88.

Four FRF are shown in Fig. F.1 of the same microphone in the four different ears of the passengers when only one loudspeaker ( $P_R$ ) is driven without any processing.

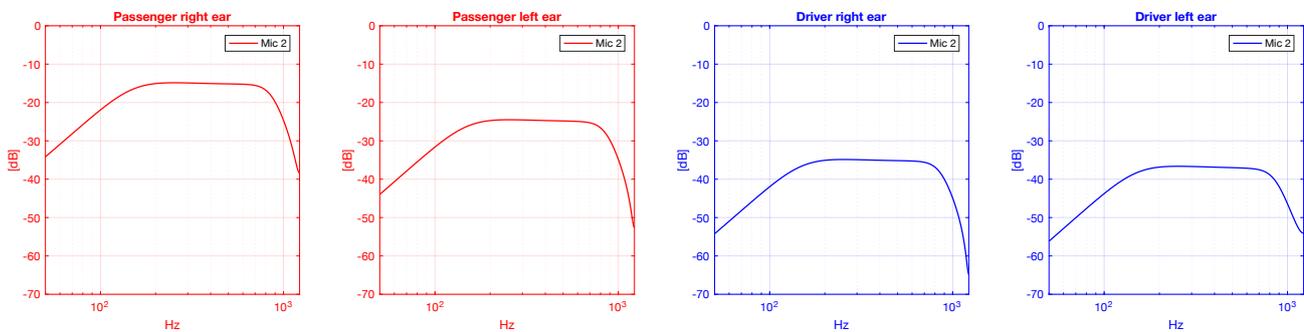


Figure F.4: FRF when only driving loudspeaker  $P_R$  without any processing for the four microphones located at the ears, (microphones 2 in Fig. 4.11).

Using Eq. 5.17 it is possible to calculate the convergence condition, for this evaluation it has been assumed that  $\tilde{\mathbf{R}} = \mathbf{R}$ , which results in  $\mu < 2.5 \times 10^{-3}$ .

In Table F.2 a summary of the definitions

<b>FxLMS parameters: preliminary study</b>	
Sampling Frequency	2450 Hz
Filter Coefficients	$M = 200$
Delay	20 samples (comprised in $\psi_d$ )
CP	Monopoles in free-field with high-pass filter
EP	Monopoles in free-field with high-pass filter
$\psi_{dBZ}$ filter (BZ)	As in the deterministic approaches = $\mathbf{p}_{dBZ}$ ( $P_L$ and $P_R$ driven in phase )
$\psi_{dDZ}$ filter (DZ)	Vector of zeros ( $\mathbf{0}$ )
Input	10 s random signal
Filters at $n = 0$	$\mathbf{w}[0] = \mathbf{0}$
Plant approx.	$\tilde{\mathbf{R}} = \mathbf{R}$
Step-size param.	$\mu = [3.5 \times 10^{-3} - 5 \times 10^{-4}]$
Mics. available	$K = 36$

Table F.2: Summary of the FxLMS parameters for the study of the  $\mu$  effects using simulated IRs.

Fig. F.5 shows the learning curve of the algorithm using simulated IRs for different values of  $\mu$ .

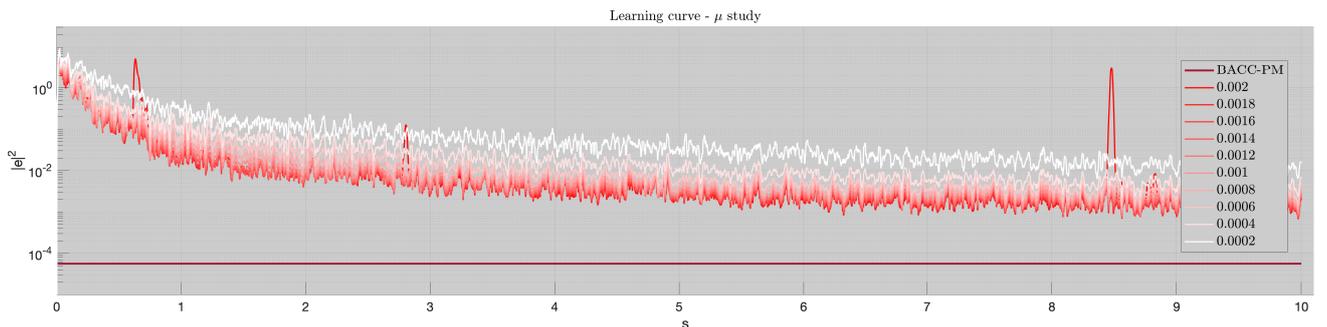


Figure F.5: Learning curve for different  $\mu$  values at Pos = 0

In order to extend these conclusions to the PSZ literature, one proceeded to plot both the resulting  $AC$  and  $E_{rr}$  using the last filter combination obtained after 10 s for different values of  $\mu$ .

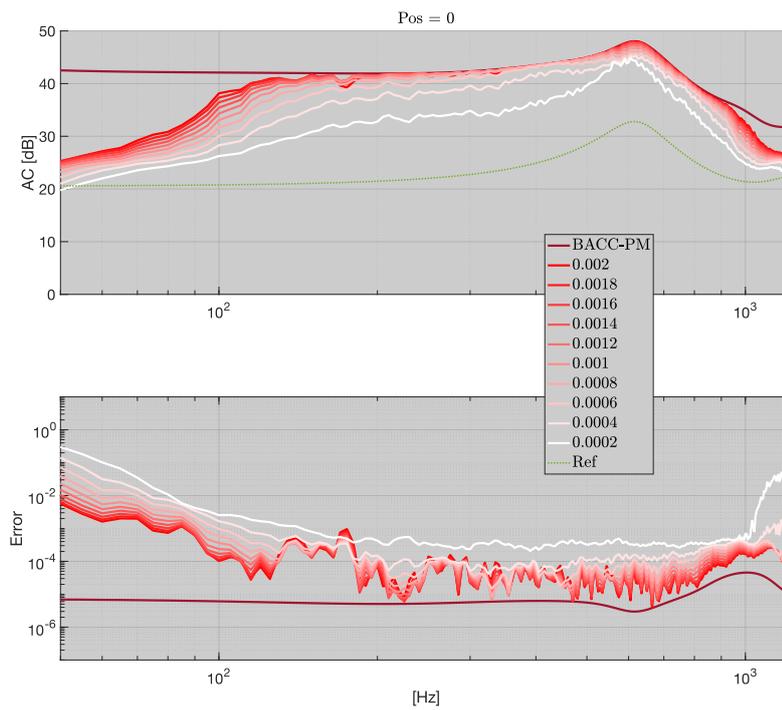


Figure F.6: Results: Metrics at Pos = 0 using FxLMS after 10 s

Even for the larger values of  $\mu$ , in Fig.F.6 the frequency dependent characteristics of the FxLMS observed in Sec. 5.3.1 are present when the band-pass frequency response is incorporated to the IRs.



# Bibliography

- [1] W. F. Druyvesteyn and J. Garas, "Personal sound," J. Audio Eng. Soc., vol. 45, no. 9, pp. 685–701, nov 1997.
- [2] J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," J. Acoust. Soc. Am., vol. 111, no. 4, pp. 1695–1700, apr 2002.
- [3] S. J. Elliott and M. Jones, "An active headrest for personal audio." J. Acoust. Soc. Am., vol. 119, no. 5 Pt 1, pp. 2702–9, may 2006.
- [4] J. Droppo, I. Tashev, and M. Seltzer, "Personal Audio Space," 2007. [Online]. Available: <https://www.microsoft.com/en-us/research/blog/personal-audio-space-headphones-experience-sans-headphones/>
- [5] J.-H. Chang, J.-y. Park, and Y.-h. Kim, "Scattering effect on the sound focused personal audio system," J. Acoust. Soc. Am., vol. 125, no. 5, pp. 3060 – 6, 2009.
- [6] S. Widmark, "Causal MSE-Optimal Filters for Personal Audio Subject to Constrained Contrast," IEEE/ACM Trans. Audio, Speech, Language Process., vol. 27, no. 5, pp. 972–987, 2019.
- [7] M. Poletti, "An investigation of 2D multizone Surround Sound Systems," 125th AES Convention, pp. 1–9, 2008.
- [8] Harman, "Individual Sound Zones," 2016. [Online]. Available: <http://www.individualsoundzones.comhttp://www.individualsoundzones.com/>
- [9] Hyundai, "Separated Sound Zone technology," 2018. [Online]. Available: <https://newatlas.com/hyundai-separate-sound-zone/55866/>
- [10] J.-y. Park, J.-h. Chang, and Y.-h. Kim, "Generation of Independent Bright Zones for a Two-Channel Private Audio System \*," J. Audio Eng. Soc., vol. 58, no. 5, 2010.
- [11] M. F. Simon, S. J. Elliott, and J. Cheer, "Loudspeaker arrays for family TV," ICSV19, pp. 1 – 8, 2012.
- [12] J.-M. Lee, T.-W. Lee, J.-Y. Park, and Y.-H. Kim, "Generation of a Private Listening Zone; Acoustic Parasol," in ICA 20, 2010, pp. 1–7.
- [13] M. B. Moller and J. Ostergaard, "A Moving Horizon Framework for Sound Zones," IEEE/ACM Trans. Audio, Speech, and Lan. Process., vol. 28, pp. 256–265, 2020.
- [14] J. Cheer, "Active control of the acoustic environment in an automobile cabin," Ph.D. dissertation, Southampton University, 2012.
- [15] J. Cheer and S. Elliott, "Design and implementation of a personal audio system in a car cabin," in Proceedings of Meetings on Acoustics, vol. 19, no. 055009, 2013, pp. 1 – 9.
- [16] J. Cheer, S. J. Elliott, M. F. S. Gálvez, and M. F. S. I. M. On, "Design and Implementation of a Car Cabin Personal Audio System," J. Audio Eng. Soc., vol. 61, no. 6, pp. 412–424, may 2013.
- [17] N. Yanagidate, J. Cheer, S. Elliott, and T. Toi, "Car cabin personal audio: Acoustic contrast with limited sound differences," in Proceedings of the 55th AES International conference: Spatial Audio, 2014, pp. 1–8.
- [18] X. Liao, S. Elliott, and J. Cheer, "Design array of loudspeakers for personal audio in a car cabin system," ICSV23, pp. 1 – 8, 2016.
- [19] X. Liao, J. Cheer, S. Elliott, and S. Zheng, "Design of a Loudspeaker Array for Personal Audio in a Car Cabin," J. Audio Eng. Soc., vol. 65, no. 3, pp. 226–238, mar 2017.
- [20] C. House, S. Dennison, D. G. Morgan, N. Rushton, G. V. White, J. Cheer, and S. Elliott, "Personal Spatial Audio in Cars Development of a loudspeaker array for multi-listener transaural reproduction in a vehicle," in Proceedings of the Institute of Acoustics, vol. 39, 2017.
- [21] S. Widmark, "Causal MMSE Filters for Personal Audio A Polynomial Matrix Approach," Ph.D. dissertation, Uppsala Universitet, 2018.
- [22] M. B. Møller and M. Olsen, "On in situ beamforming in an automotive cabin using a planar loudspeaker array," in PROCEEDINGS of the 23rd International Congress on Acoustics, 2019, pp. 1109–1116.

- [23] J.-h. Chang and W.-h. Cho, "Evaluation of independent sound zones in a car," in PROCEEDINGS of the 23rd International Congress on Acoustics, 2019, pp. 5196–5203.
- [24] M. Christoph and M. Kronlachner, "Improvement of Personal Sound Zones by Individual Delay Compensation," AES Conference on Sound Field Control, pp. 1 – 8, 2016.
- [25] E. LaCarrubba, "Acoustic reproduction device with improved directional characteristics," United States Patent, no. US 6,820,718 B2, pp. 1 – 10, 2004.
- [26] M. Olsen and M. B. Møller, "Sound zones: on the effect of ambient temperature variations in feed-forward systems," AES 142nd Convention, pp. 1 – 10, 2017.
- [27] H. So and J.-W. Choi, "Subband Optimization and Filtering Technique," ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, pp. 8494–8498, 2019.
- [28] M. Jones and S. J. Elliott, "Personal audio with multiple dark zones." J. Acoust. Soc. Am, vol. 124, no. September, pp. 3497–3506, 2008.
- [29] J. Francombe, R. Mason, M. Dewhurst, and S. Bech, "Determining the Threshold of Acceptability for an Interfering Audio Programme," AES 132nd Convention, pp. 1 – 17, 2012.
- [30] K. Baykaner, P. Coleman, R. Mason, P. Jackson, J. Francombe, M. Olik, and S. Bech, "The Relationship Between Target Quality and Interference in Sound Zone," J. Audio Eng. Soc., vol. 63, no. 1/2, pp. 78–89, jan 2015.
- [31] Sennheiser, "AMBEO Mobility Audio," 2021. [Online]. Available: <https://fr-fr.sennheiser.com/ambeo-mobility>
- [32] Devialet, "Symbioz demo car," 2021. [Online]. Available: <https://www.deviallet.com/en-eu/about-us/our-stories/story-symbioz-demo-car/>
- [33] T. Betlehem, W. Zhang, M. A. Poletti, and T. D. Abhayapala, "Personal sound zones: Delivering interface-free audio to multiple listeners," IEEE Signal Processing Magazine, vol. 32, no. 2, pp. 81–91, 2015.
- [34] J. Cheer, S. J. Elliott, Y. Kim, and J.-w. Choi, "Practical Implementation of Personal Audio in a Mobile Device\*," J. Audio Eng. Soc, vol. 61, no. 5, pp. 290–300, 2013.
- [35] D. A. Russell, J. P. Titlow, and Y.-J. Bommen, "Acoustic monopoles, dipoles, and quadrupoles: An experiment revisited," American Journal of Physics, vol. 67, no. 8, pp. 660–664, 1999.
- [36] L. Beranek, Acoustics, 1993rd ed. Cambridge: Acoustical Society of America, 1993.
- [37] C. a. Balanis, Antenna Theory and Design, 3rd ed. John Wiley and Sons, 2005, no. 3.
- [38] K. G. Foote, "Discriminating between the nearfield and the farfield of acoustic transducers," J. Acoust. Soc. Am, vol. 136, no. 4, pp. 1511–1517, 2014.
- [39] M. M. Boone and O. Ouweltjes, "Design of a loudspeaker system with a low-frequency cardioidlike radiation pattern," Journal of the Audio Engineering Society, vol. 45, no. 9, pp. 702–707, 1997.
- [40] H. F. Olson, "Gradient Loudspeakers," J. Audio Eng. Soc., vol. 21, no. 2, pp. 86–93, jan 1973.
- [41] J. Kuutti, J. Leiwo, and R. Sepponen, "Local Control of Audio Environment: A Review of Methods and Applications," Technologies, vol. 2, no. 1, pp. 31–53, 2014.
- [42] F. Zangeneh-Nejad and R. Fleury, "Active times for acoustic metamaterials," Reviews in Physics, vol. 4, no. April, pp. 1 – 17, 2019.
- [43] Y. Cheng and X. Liu, "Coupled resonant modes in twisted acoustic metamaterials," Applied Physics A: Materials Science and Processing, vol. 109, no. 4, pp. 805–811, 2012.
- [44] X. Liu and J. Liu, "Control of sound directivity based on metamaterials," in INTER-NOISE 2018 - 47th International Congress and Exposition on Noise Control Engineering: Impact of Noise Control Engineering, Chicago, 2018, pp. 1–8.
- [45] Y. Cheng, J. Y. Xu, and X. J. Liu, "Tunable sound directional beaming assisted by acoustic surface wave," Applied Physics Letters, vol. 96, no. 7, pp. 1–4, 2010.
- [46] Sennheiser and Continental, "Speakerless Immersive Sound," 2020. [Online]. Available: <https://www.continental.com/en/press/press-releases/2020-01-06-ces-206016>
- [47] Faurecia, "Immersive experiences that enhance time spent on the move." [Online]. Available: <https://www.faurecia-clarion.com/technology/immersive-experiences>
- [48] O. Jeon, H. Ryu, and S. Wang, "Optimization of Exciter Arrangement to Improve Beamforming Performance of Multi-Actuator Panels with Low-Damping Loss Factor," in INTER-NOISE 2018 - 47th International Congress and Exposition on Noise Control Engineering: Impact of Noise Control Engineering, Chicago, 2018, pp. 1–5.
- [49] O. Jeon, H. Ryu, and W. Semyung, "Optimal design of exciter array to minimize bending mode effect of multi-actuator panels in beamforming," INTER-NOISE 2019 MADRID - 48th International Congress and Exhibition on Noise Control Engineering, pp. 3–7, 2019.

- [50] T. J. Holmes, "The "Acoustic Resistance Box" - A Fresh Look at an Old Principle," *J. Audio Eng. Soc.*, vol. 34, no. 12, pp. 981–989, 1986.
- [51] D. Corynen, "Dipole loudspeaker for reproducing sound at bass frequencies," *International Patent*, no. WO 2019/121266, pp. 1–22, 2019.
- [52] D. Corynen, "Dipole loudspeaker for producing sound at bass frequencies," *International Patent*, no. WO 2020/234317, pp. 1–24, 2020.
- [53] T. Mellow and L. Kärkkäinen, "On the sound field of an oscillating disk in a finite open and closed circular baffle," *J. Acoust. Soc. Am.*, vol. 118, no. 3, pp. 1311–1325, 2005.
- [54] M. Van Der Wal, W. S. Evert, and D. V. Diemer, "Design of Logarithmically Spaced Constant-Directivity Transducer Arrays," *J. Audio Eng. Soc.*, vol. 44, no. 6, pp. 497–507, 1996.
- [55] I. Wolff and L. Malter, "Directional Radiation of Sound," *J. Acoust. Soc. Am.*, vol. 2, no. 2, pp. 201–241, oct 1930.
- [56] D. Havelock and A. Brammer, "Directional loudspeakers using sound beams," *J. Audio Eng. Soc.*, vol. 48, no. 10, pp. 908–916, 2000.
- [57] K. R. Holland and F. J. Fahy, "A Low-Cost End-Fire Acoustic Radiator\*," *J. Audio Eng. Soc.*, vol. 39, no. 7, pp. 540–550, 1991.
- [58] B. Van Veen and K. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, apr 1988.
- [59] D. G. Meyer, "Multiple-Beam, Electronically Steered Line-Source Arrays for Sound-Reinforcement Applications\*," *J. Audio Eng. Soc.*, vol. 38, no. 4, pp. 237–249, 1990.
- [60] M. S. Ureda, "Analysis of loudspeaker line arrays," *J. Audio Eng. Soc.*, vol. 52, no. 5, pp. 467–495, 2004.
- [61] S. Palk, M. De Geest, and K. Vansant, "Interior acoustic simulation for in-car audio design," *Sound and Vibration*, vol. 47, no. 1, pp. 10–17, 2013.
- [62] A. Putra, A. L. Munawir, and W. M. Farid, "Corrected Statistical Energy Analysis Model for Car Interior Noise," *Advances in Mechanical Engineering*, vol. 7, no. 1, 2015.
- [63] P. J. Westervelt, "Parametric Acoustic Array," *J. Acoust. Soc. Am.*, vol. 35, no. 4, pp. 535–537, apr 1963.
- [64] M. Yoneyama, J.-I. Fujimoto, Y. Kawamo, and S. Sasabe, "The audio spotlight: An application of nonlinear interaction of sound waves to a new type of loudspeaker design," *J. Acoust. Soc. Am.*, vol. 73, no. 5, pp. 1532–1536, 1983.
- [65] Soundlazer, "Soundlazer." [Online]. Available: <https://www.soundlazer.com>
- [66] C. Shi, H. Nomura, T. Kamakura, and W.-S. Gan, "Development of a steerable stereophonic parametric loudspeaker," in *Signal and Information Processing Association Annual Summit and Conference (APSIPA)*. IEEE, dec 2014, pp. 1–5.
- [67] W.-S. Gan, J. Yang, and T. Kamakura, "A review of parametric acoustic array in air," *Applied Acoustics*, vol. 73, no. 12, pp. 1211–1219, dec 2012.
- [68] J. Ahrens, *Analytic Methods of Sound Field Synthesis*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.
- [69] W. Jin, "Spatial multizone soundfield reproduction design," Ph.D. dissertation, Victoria University of Wellington, 2015.
- [70] Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 6, pp. 1711–1720, 2011.
- [71] S. Spors and J. Ahrens, "A Comparison of Wave Field Synthesis and Higher-Order Ambisonics with Respect to Physical Properties and Spatial Sampling," *AES 125th Convention*, no. June 2014, pp. 1–17, 2008.
- [72] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, may 1993.
- [73] J. Daniel, R. Nicol, and S. Moreau, "Further investigations of high order ambisonics and wavefield synthesis for holo-phonic sound imaging," *AES 114th Convention*, pp. 1 – 18, 2003.
- [74] Y. J. Wu and T. D. Abhayapala, "Theory and design of soundfield reproduction using continuous loudspeaker concept," *IEEE Trans. Audio, Speech, Language Process.*, vol. 17, no. 1, pp. 107–116, 2009.
- [75] S. Spors, "Active Listening Room Compensation for Spatial Audio Reproduction Systems," Ph.D. dissertation, Der Technischen Fakultät der Friedrich-Alexander-Universität Erlangen-Nürnberg, 2005.
- [76] E. Verheijen, "Sound Reproduction by Wave Field Synthesis," Ph.D. dissertation, Delft University of Technology, 1997.
- [77] P.-A. Gauthier and A. Berry, "Adaptive wave field synthesis with independent radiation mode control for active sound field reproduction: Theory," *J. Acoust. Soc. Am.*, vol. 119, no. 5, pp. 2721–2737, 2006.
- [78] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter, "Spatial Sound With Loudspeakers and Its Perception: A Review of the Current State," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 1920–1938, sep 2013.

- [79] L. Romoli, P. Peretti, L. Palestini, S. Cecchi, and F. Piazza, "A new approach to digital directivity control of loudspeakers line arrays using wave field synthesis theory," Proc. Int. Workshop on Acoustic Signal Enhancement (IWAENC), pp. 2–5, 2008.
- [80] T. Okamoto, "Generation of multiple sound zones by spatial filtering in wavenumber domain using a linear array of loudspeakers," ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, pp. 4733–4737, 2014.
- [81] H. Itou, K. Furuya, and Y. Haneda, "Evanescent wave reproduction using linear array of loudspeakers," in 2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), no. 3. IEEE, oct 2011, pp. 37–40.
- [82] H. Itou, K. Furuya, and Y. Haneda, "Localized sound reproduction using circular loudspeaker array based on acoustic evanescent wave," in 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), no. February. IEEE, mar 2012, pp. 221–224.
- [83] P.-A. Gauthier and A. Berry, "Adaptive wave field synthesis for broadband active sound field reproduction: Signal processing," J. Acoust. Soc. Am., vol. 123, no. 4, pp. 2003–2016, 2008.
- [84] S. Spors, H. Buchner, R. Rabenstein, and W. Herboldt, "Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering," J. Acoust. Soc. Am., vol. 122, no. 1, pp. 354–369, 2007.
- [85] M. Schneider and W. Kellermann, "Multichannel acoustic echo cancellation in the wave domain with increased robustness to nonuniqueness," IEEE/ACM Trans. Audio, Speech, Language Process., vol. 24, no. 3, pp. 518–529, 2016.
- [86] M. Olik, "Personal Sound Zone Reproduction with Room Reflections," Ph.D. dissertation, University of Surrey, 2014.
- [87] M. Olsen and M. B. Møller, "Sound zones : scattering study with head and torso simulator," Proceedings of the 52nd international conference of Audio Engineering Society, pp. 1–9, 2013.
- [88] F. Jacobsen, M. Olsen, M. Moller, and F. Agerkvist, "A comparison of two strategies for generating sound zones in a room," ICSV18, no. July, pp. 1 – 8, 2011.
- [89] W. Zhang, T. D. Abhayapala, T. Betlehem, and F. M. Fazi, "Analysis and control of multi-zone sound field reproduction using modal-domain approach," J. Acoust. Soc. Am., vol. 140, no. 3, pp. 2134–2144, sep 2016.
- [90] W. Jin, W. B. Kleijn, and D. Virette, "Multizone soundfield reproduction using orthogonal basis expansion," ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, pp. 311–315, 2013.
- [91] J. Donley, C. Ritz, and W. B. Kleijn, "Multizone Soundfield Reproduction with Privacy- and Quality-Based Speech Masking Filters," IEEE/ACM Trans. Audio, Speech, Language Process., vol. 26, no. 6, pp. 1037–1051, 2018.
- [92] M. a. Poletti and F. M. Fazi, "An approach to generating two zones of silence with application to personal sound systems," J. Acoust. Soc. Am., vol. 137, no. 2, pp. 598–605, feb 2015.
- [93] Audi, "New approaches to perfect sound – Hi-fi expertise at Audi," p. 1, 2010. [Online]. Available: <https://www.eurocarnews.com/57/0/696/0/new-approaches-to-perfect-sound-hi-fi-expertise-at-audi.html>
- [94] M. B. Moller, "Sound Zone Control inside Spatially Confined Regions in Acoustic Enclosures," Ph.D. dissertation, Aalborg University, 2019.
- [95] M. Buerger, C. Hofmann, and W. Kellermann, "Broadband Multizone Sound Rendering by Jointly Optimizing the Sound Pressure and Particle Velocity," J. Acoust. Soc. Am., vol. 143, no. 3, pp. 1477–1490, mar 2018.
- [96] P. Coleman, "Loudspeaker array processing for personal sound zone reproduction," Ph.D. dissertation, Surrey University, 2014.
- [97] M. F. Simón Gálvez, S. J. Elliott, and J. Cheer, "Time Domain Optimization of Filters Used in a Loudspeaker Array for Personal Audio," IEEE/ACM Trans. on Audio, Speech, and Lan. Process., vol. 23, no. 11, pp. 1869–1878, nov 2015.
- [98] S. J. Elliott, J. Cheer, H. Murfet, and K. R. Holland, "Minimally radiating sources for personal audio," J. Acoust. Soc. Am., vol. 128, no. 4, pp. 1721–1728, 2010.
- [99] M. F. Simón Gálvez, S. J. Elliott, and J. Cheer, "A superdirective array of phase shift sources," J. Acoust. Soc. Am., vol. 132, no. 2, pp. 746–756, aug 2012.
- [100] M. F. S. Gálvez, S. J. Elliott, and J. Cheer, "Superdirective beamforming in rooms," Revista de acústica, vol. 45, pp. 17–24, 2013.
- [101] P. J. Jackson, F. Jacobsen, P. Coleman, and J. Abildgaard Pedersen, "Sound field planarity characterized by superdirective beamforming," in ICA, vol. 19, no. 055056, 2013, pp. 1 – 9.
- [102] F. Olivieri, F. M. Fazi, M. Shin, and P. Nelson, "Pressure-Matching beamforming method for loudspeaker arrays with frequency dependent selection of control points," AES 138th Convention, pp. 1–10, 2015.
- [103] X. Ma, C. Hohnerlein, and J. Ahrens, "Listener-position adaptive crosstalk cancelation using a parameterized superdirective beamformer," Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop, pp. 114–118, 2018.

- [104] Z. Tu, J. Lu, and X. Qiu, "Robustness of a compact endfire personal audio system against scattering effects (L)," *J. Acoust. Soc. Am.*, vol. 140, no. 4, pp. 2720–2724, 2016.
- [105] Y. J. Wu and T. D. Abhayapala, "Multizone soundfield reproduction using multiple loudspeaker arrays," *Proceedings of 20th International Congress on Acoustics*, pp. 1 – 5, 2010.
- [106] M. B. Møller and M. Olsen, "Sound Zones: On Performance Prediction of Contrast Control Methods," in *2016 AES International Conference on Sound Field Control*, no. 95, 2016, pp. 1–10.
- [107] M. Sanalati, "Synthèse d'un champ acoustique avec contraste spatial élevé," Ph.D. dissertation, Le Mans Universite, 2018.
- [108] M. Bouchard and S. Quednau, "Multichannel recursive-least-squares algorithms and fast-transversal-filter algorithms for active noise control and sound reproduction systems," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 606–618, 2000.
- [109] M. Buerger, C. Hofmann, C. Frankenbach, and W. Kellermann, "Multizone sound reproduction in reverberant environments using an iterative least-squares filter design method with a spatiotemporal weighting function," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, oct 2017, pp. 1–5.
- [110] M. Schneider and E. A. Habets, "Iterative DFT-Domain Inverse Filter Optimization Using a Weighted Least-Squares Criterion," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 27, no. 12, pp. 1957–1969, 2019.
- [111] G. Roussel, "Contributions à la mise au point de méthodes adaptatives de reproduction de champs sonores multi-zone pour les auditeurs en mouvement," Ph.D. dissertation, Le Mans Universite, 2019.
- [112] Y. Cai, M. Wu, L. Liu, and J. Yang, "Time-domain acoustic contrast control design with response differential constraint in personal audio systems," *J. Acoust. Soc. Am.*, vol. 135, no. 6, pp. EL252–EL257, 2014.
- [113] D. H. M. Schellekens, M. B. Moller, and M. Olsen, "Time domain acoustic contrast control implementation of sound zones for low-frequency input signals," *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 365–369, mar 2016.
- [114] S. Berthilsson, A. Barkefors, L.-j. Brannmark, and M. Sternad, "Acoustical Zone Reproduction for Car Interiors Using a MIMO MSE Framework Acoustical Zone Reproduction for Car Interiors Using a MIMO MSE Framework," *AES 48th International Conference*, no. September, 2012.
- [115] P. Coleman, P. J. B. Jackson, M. Olik, M. Møller, M. Olsen, and J. Abildgaard Pedersen, "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array," *J. Acoust. Soc. Am.*, vol. 135, no. 4, pp. 1929–1940, 2014.
- [116] S. J. Elliott and J. Cheer, "ISVR Technical Memorandum: Regularisation and Robustness of Personal Audio Systems," Southampton University, dec 2011.
- [117] A. D. Belegundu, T. R. C, and Chandrupatla, *Optimization Concepts and Applications in Engineering*, 2nd ed. Cambridge University Press, 2011.
- [118] M. B. Møller and M. Olsen, "Sound Zones : on Envelope Shaping of Fir Filters," *ICSV24*, pp. 613–620, 2017.
- [119] M. Shin, S. Q. Lee, F. M. Fazi, P. A. Nelson, D. Kim, S. Wang, K. H. Park, and J. Seo, "Maximization of acoustic energy difference between two spaces," *J. Acoust. Soc. Am.*, vol. 128, no. 1, pp. 121–131, 2010.
- [120] P. Coleman, P. Jackson, M. Olik, and J. A. Pedersen, "Optimizing the planarity of sound zones," *Proceedings of the 52nd international conference of Audio Engineering Society*, pp. 1–10, 2013.
- [121] Y. Cai, M. Wu, and J. Yang, "Sound reproduction in personal audio systems using the least-squares approach with acoustic contrast control constraint," *J. Acoust. Soc. Am.*, vol. 135, no. 2, pp. 734–741, 2014.
- [122] Y. Cai, M. Wu, and J. Yang, "Design of a time-domain acoustic contrast control for broadband input signals in personal audio systems," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, may 2013, pp. 341–345.
- [123] O. Kirkeby and P. A. Nelson, "Reproduction of plane wave sound fields," *J. Acoust. Soc. Am.*, vol. 94, no. 5, pp. 2992–3000, nov 1993.
- [124] S. Haykin, *Adaptive Filter Theory Fifth Edition*, 5th ed. London: Pearson, 2014.
- [125] L. Vindrola, M. Melon, J.-C. Chamard, B. Gazengel, and G. Plantier, "Personal Sound Zones: a comparison between frequency and time domain formulations in a transportation context," *AES 147 Convention*, 2019.
- [126] J.-H. Chang and F. Jacobsen, "Sound field control with a circular double-layer array of loudspeakers," *J. Acoust. Soc. Am.*, vol. 131, no. 6, pp. 4518–4525, 2012.
- [127] V. Moles-Cases, G. Pinero, M. De Diego, and A. Gonzalez, "Personal Sound Zones by Subband Filtering and Time Domain Optimization," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 28, pp. 2684–2696, 2020.
- [128] M. Bo Moller, J. Kjaer Nielsen, E. Fernandez-Grande, and S. Krarup Olesen, "On the Influence of Transfer Function Noise on Low Frequency Pressure Matching for Sound Zones," in *IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, jul 2018, pp. 331–335.

- [129] S. J. Elliott, J. Cheer, J.-W. Choi, and Y. Kim, "Robustness and Regularization of Personal Audio Systems," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 7, pp. 2123–2133, sep 2012.
- [130] P.-A. Gauthier, Y. Pasco, and A. Berry, "Generalized singular value decomposition for personalized audio using loudspeaker array," in *AES International Conference on Sound Field Control*, 2016, pp. 1–10.
- [131] Q. Zhu, P. Coleman, X. Qiu, M. Wu, J. Yang, and I. Burnett, "Robust Personal Audio Geometry Optimization in the SVD-Based Modal Domain," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 27, no. 3, pp. 610–620, 2019.
- [132] T. Lee, J. K. Nielsen, and M. G. Christensen, "Signal-Adaptive and Perceptually Optimized Sound Zones with Variable Span Trade-Off Filters," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 28, no. 8, pp. 2412–2426, 2020.
- [133] T. Lee, J. K. Nielsen, J. R. Jensen, and M. G. Christensen, "A unified approach to generating sound zones using variable span linear filters," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2018-April, no. April, pp. 491–495, 2018.
- [134] J. K. Nielsen, T. Lee, R. J. Jensen, and M. G. Christensen, "Sound zones as an optimal filtering problem," *52nd Asilomar Conference*, 2018.
- [135] J. Ramo, S. Marsh, S. Bech, R. Mason, and S. Holdt Jenssen, "Validation of a perceptual distraction model in a complex personal sound zone system," in *AES 141st Convention*, 2016, pp. 1–10.
- [136] J. Francombe, R. Mason, M. Dewhurst, and S. Bech, "Elicitation of attributes for the evaluation of audio-on-audio interference," *The Journal of the Acoustical Society of America*, vol. 136, no. 5, pp. 2630–2641, 2014.
- [137] J. Francombe, "Perceptual evaluation of audio-on-audio Interference in a personal sound zone system," Ph.D. dissertation, Surrey University, 2014.
- [138] K. R. Baykaner, "Predicting the perceptual acceptability of auditory interference for the optimisation of sound zones," Ph.D. dissertation, University of Surrey, 2014.
- [139] D. Wallace and J. Cheer, "Optimisation of personal audio systems for intelligibility contrast," *144th AES Convention*, pp. 1 – 10, 2018.
- [140] P. Coleman, P. J. B. Jackson, M. Olik, and J. A. Pedersen, "Numerical optimization of loudspeaker configuration for sound zone reproduction," *ICSV21*, pp. 1 – 8, 2014.
- [141] J. Vanderkooy, "The Acoustic Center: A New Concept for Loudspeakers at Low Frequencies," in *121st AES Convention*, 2006, pp. 1 – 15.
- [142] X. Ma, P. J. Hegarty, and J. J. Larsen, "Mitigation of Nonlinear Distortion in Sound Zone Control by Constraining Individual Loudspeaker Driver Amplitudes," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 456–460, 2018.
- [143] H. Møller, "Fundamentals of binaural technology," *Applied Acoustics*, vol. 36, no. 3-4, pp. 171–218, 1992.
- [144] P. A. Nelson, H. Hamada, and S. J. Elliott, "Adaptive Inverse Filters for Stereophonic Sound Reproduction," *IEEE Trans. Sig. Process.*, vol. 40, no. 7, pp. 1621–1632, 1992.
- [145] P. Coleman, P. J. Jackson, M. Olik, and J. A. Pedersen, "Stereophonic personal audio reproduction using planarity control optimization," *ICSV21*, pp. 1 – 8, 2014.
- [146] A. Vidal, "Diffusion de son 3D par synthèse de champs acoustiques binauraux," Ph.D. dissertation, Université d'Aix-Marseille, 2017.
- [147] T. Lee, L. Shi, J. K. Nielsen, and M. G. Christensen, "Fast Generation of Sound Zones Using Variable Span Trade-Off Filters in the DFT-Domain," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 29, pp. 363–378, 2021.
- [148] L. Shi, T. Lee, L. Zhang, J. K. Nielsen, and M. G. Christensen, "A fast reduced-rank sound zone control algorithm using the conjugate gradient method," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2020-May, pp. 436–440, 2020.
- [149] J.-Y. Park, J.-W. Choi, and Y.-H. Kim, "Acoustic contrast sensitivity to transfer function errors in the design of a personal audio system," *J. Acoust. Soc. Am.*, vol. 134, no. 1, pp. EL112–EL118, 2013.
- [150] Q. Zhu, P. Coleman, M. Wu, and J. Yang, "Robust acoustic contrast control with reduced in-situ measurement by acoustic modeling," *J. Audio Eng. Soc.*, vol. 65, no. 6, pp. 460–473, 2017.
- [151] X. Ma, P. J. Hegarty, and J. A. Pedersen, "Assessing the influence of loudspeaker driver nonlinear distortion on personal sound zones," *AES 142nd Convention*, pp. 1 – 7, 2017.
- [152] W.-h. Cho and J.-h. Chang, "Investigation of effect on the acoustic transfer function in a vehicle cabin according to change of configuration," in *PROCEEDINGS of the 23rd International Congress on Acoustics*, 2019, pp. 5212–5217.
- [153] M. F. Simón Gálvez and F. M. Fazi, "Listener adaptive filtering strategies for personal audio reproduction over loudspeaker arrays," in *AES Conference on Sound Field Control*, jul 2016, pp. 1 – 9.
- [154] M. A. Sanderson and T. Onsay, "CAE interior cavity model validation using acoustic modal analysis," *SAE Technical Papers*, pp. 1 – 8, 2007.

- [155] S. H. Kim, J. M. Lee, and M. H. Sung, "Structural-acoustic modal coupling analysis and application to noise reduction in a vehicle passenger compartment," *J. Sound Vib.*, vol. 225, no. 5, pp. 989–999, 1999.
- [156] Y. Haneda, Y. Kaneda, and N. Kitawaki, "Common-Acoustical-Pole and Residue Model and Its Application to Spatial Interpolation and Extrapolation of a Room Transfer Function," *IEEE Speech Audio Process.*, vol. 7, no. 6, pp. 709–717, 1999.
- [157] D. Moreau, B. Cazzolato, A. Zander, and C. Petersen, "A review of virtual sensing algorithms for active noise control," *Algorithms*, vol. 1, no. 2, pp. 69–99, 2008.
- [158] A. Roure and A. Albarrazin, "The remote microphone technique for active noise control," in *Active control of sound and vibration; Proceedings of ACTIVE 99*, 1999, pp. 1233–1244.
- [159] W. Jung, S. J. Elliott, and J. Cheer, "Combining the remote microphone technique with head-tracking for local active sound control," *J. Acoust. Soc. Am.*, vol. 142, no. 1, pp. 298–307, 2017.
- [160] W. Jung, S. J. Elliott, and J. Cheer, "Estimation of the pressure at a listener's ears in an active headrest system using the remote microphone technique," *J. Acoust. Soc. Am.*, vol. 143, no. 5, pp. 2858–2869, 2018.
- [161] D. P. Das, D. J. Moreau, and B. Cazzolato, "Performance evaluation of an active headrest using the remote microphone technique," *Proceedings of the Annual Conference on the Australian Acoustical Society*, no. 69, pp. 1 – 8, 2011.
- [162] A. Benichoux, E. Vincent, and R. Gribonval, "A compressed sensing approach to the simultaneous recording of multiple room impulse responses," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 285–288, 2011.
- [163] L. Vindrola, M. Melon, J.-C. Chamard, and B. Gazengel, "Pressure Matching With Forced Filters for Personal Sound Zones Application," *J. Audio Eng. Soc.*, vol. 68, no. 11, pp. 832–842, dec 2020.
- [164] B. Widrow and M. Hoff, "Adaptive Switching Circuits," New York, NY, pp. 96 – 104, 1960.
- [165] G. A. Clark, S. R. Parker, and S. K. Mitra, "Block Implementation of Adaptive Digital Filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 3, pp. 744–752, 1981.
- [166] D. P. Das, G. Panda, and S. M. Kuo, "New block filtered-x LMS algorithms for active noise control systems," *IET Signal Process.*, vol. 1, no. 2, pp. 73–81, 2007.
- [167] S. Elliott, *Signal Processing for Active control*. London: Academic, 2001.
- [168] C. Antoñanzas, M. Ferrer, M. De Diego, and A. Gonzalez, "Personal active control systems over coupled networks," *24th International Congress on Sound and Vibration, ICSV 2017*, no. July, 2017.
- [169] J. K. Thomas, S. P. Lovstedt, J. D. Blotter, and S. D. Sommerfeldt, "Eigenvalue equalization filtered-x algorithm for the multichannel active noise control of stationary and nonstationary signals," *J. Acoust. Soc. Am.*, vol. 123, no. 6, pp. 4238–4249, 2008.
- [170] P. S. Diniz, M. L. de Campos, and A. Antoniou, "Analysis of LMS-Newton Adaptive Filtering Algorithms with Variable Convergence Factor," *IEEE Trans. Signal Process.*, vol. 43, no. 3, pp. 617–627, 1995.
- [171] T. Aboulnasr and K. Mayyas, "A robust variable step-Size lms-type algorithm: Analysis and simulations," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 631–639, 1997.



# Résumé étendu

De nos jours, les véhicules automobiles sont des moyens de transport mais également des objets de luxe et de confort. De nombreuses branches de l'ingénierie sont impliquées dans leur développement, non seulement pour améliorer la sécurité et réduire les émissions, mais aussi pour proposer une expérience originale aux passagers.

Dans le cadre des applications « audio », le nombre de systèmes de restitution sonore dans l'habitacle d'un véhicule a augmenté ces dernières années, tant pour aider à conduire que pour divertir les passagers. Le problème est que toutes ces sources sonores ne sont pas nécessairement destinées à tous les passagers, ce qui peut entraîner une distraction et/ou une gêne. Dans ce contexte, le but de cette thèse est de concevoir un système de restitution sonore capable de reproduire des programmes audio dédiés aux différents passagers d'une voiture, également connu sous le nom de bulles sonores ou zones sonores personnelles (PSZ).

La génération des PSZ est un sujet d'intérêt croissant depuis ces vingt dernières années pour un grand nombre d'applications, par exemple les téléviseurs, les avions, la sonorisation des lieux publics, les applications domestiques, etc. Néanmoins, l'application PSZ la plus étudiée est probablement l'industrie automobile.

Dès le début de la thèse, les exigences et les contraintes convenues avec le fabricant (Stellantis) sont exposées. Un résumé est détaillé dans la liste suivante :

- 1) Bande passante d'intérêt : fréquence inférieure à 1 kHz. On suppose qu'un autre système est dédié aux fréquences les plus élevées (antenne de haut-parleurs, réflecteurs acoustiques, etc.).
- 2) Zones : elles sont définies par des rectangles de 4 cm × 10 cm au niveau des oreilles du conducteur et du passager avant.
- 3) "Séparation acoustique" entre les zones : contraste acoustique d'environ 30 dB.
- 4) "Contrainte de transparence" : faible erreur de reproduction normalisée.
- 5) Nombre de canaux indépendants : moins de 10.
- 6) Faible Effort.
- 7) Habitacle automobile à propriétés acoustiques variables.

Tout d'abord, les méthodes d'optimisation sous contrainte ont été mises en œuvre en utilisant une approche déterministe. Il a été constaté que la performance du système global dépend de la définition de la fonction de coût ainsi que du problème acoustique à résoudre (définition des zones, champ acoustique cible, positions des sources, environnement acoustique, etc.). Les zones à contrôler sont définies dans le cahier des charges et correspondent aux zones autour des oreilles du conducteur (zone de silence ou Dark Zone DZ) et du passager avant (zone de signal ou Bright Zone BZ). Pour la fonction coût, il a été décidé d'utiliser un cas particulier de la méthode « pressure matching » plus connue sous le nom de « Combined Solution » qui permet de réduire au maximum l'énergie rayonnée dans la zone de silence DZ tout en reproduisant la pression souhaitée dans la zone de signal BZ.

Ce travail utilise comme point de départ des appui-têtes à deux haut-parleurs pour chacun des sièges (ici le nombre total de haut-parleur  $L$  est donc égal à 4). La pression cible dans la zone de signal BZ est la réponse naturelle de l'appui-tête de ce siège. Il a été observé que le système "classique" (appui-tête à deux haut-parleurs) était incapable de répondre aux contraintes initiales. Par conséquent, le nombre de haut-parleurs a été augmenté afin d'obtenir le contraste acoustique souhaité en détériorant le moins possible la reproduction dans la zone de signal BZ.

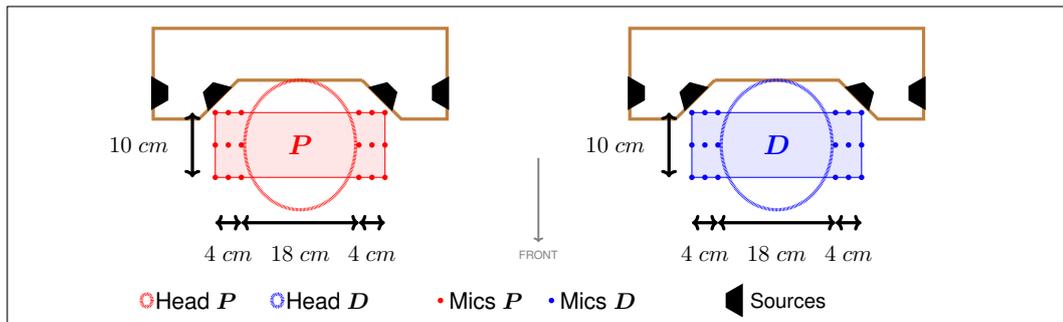


Figure R.1: Principe du système PSZ à huit sources.

Le système final est composé du système classique d'appui-tête (deux sources par siège), auquel on a ajouté des haut-parleurs supplémentaires : deux sources intérieures (entre les sièges) et deux sources extérieures, ce qui donne un total de  $L = 8$  (voir Fig. R.1). Plusieurs études (simulations, mesures) ont été mises en œuvre afin de concevoir et de tester le nouveau système de PSZ. Ces études ont été organisées en cinq niveaux de complexité acoustique, partant d'un modèle monopôles en champ libre pour aller jusqu'à l'habitacle de voiture. La figure ci-dessous montre le contraste acoustique ( $AC$ ), le conditionnement, l'effort ( $E_{ff}$ ) et l'erreur de reproduction normalisée ( $E_{rr}$ ) qui résultent de l'utilisation du système final dans l'habitacle d'une voiture. Celui-ci a permis d'obtenir un contraste ( $AC$ ) moyen de 34,7 dB et une erreur moyenne de  $1.69 \times 10^{-4}$  sur la largeur de bande étudiée dans l'environnement le plus complexe qu'est l'habitacle de voiture (voir Fig. R.2).

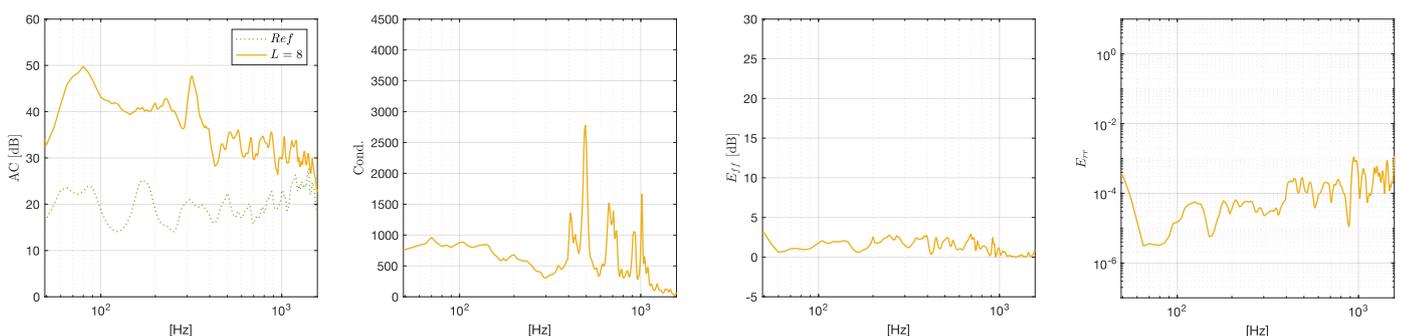


Figure R.2: Contraste Acoustique ( $AC$ ), conditionnement, erreur et effort obtenus avec le système à huit haut-parleurs dans un habitacle automobile. Pour le Contraste Acoustique ( $AC$ ), la référence ( $Ref$ ) correspond au cas sans traitement.

Les difficultés principales rencontrées lors de l'utilisation des systèmes basés sur l'optimisation sous contrainte (CO) sont le conditionnement (avoir un problème mal conditionné) et l'effort (demander au système un niveau qu'il ne peut pas produire). L'approche consistant à utiliser des sources en champ proche, à optimiser les positions des sources ajoutées, ainsi que la définition de la pression cible (la réponse naturelle du système) ont abouti à un problème "bien posé sur le plan acoustique". En d'autres termes,

dans toutes les expériences réalisées dans le cadre de cette thèse, l'effort ( $E_{ff}$ ) a été maintenu en dessous de 5 dB et le conditionnement n'a présenté aucun problème particulier (voir Fig. R.2).

Les autres difficultés concernent l'allure temporelle des filtres résultants (leur causalité, la présence du « pre-ringing », etc.). En raison du positionnement des haut-parleurs au niveau de la zone de signal BZ (en champ proche) et de la définition de la pression cible, les filtres qui contribuent le plus à la reproduction sont des impulsions de Dirac ( $\delta[n]$ ) presque parfaites. Cela permet de supposer que la distorsion (dans la reproduction) due au traitement PSZ est négligeable.

D'autre part, deux formulations différentes de la méthode "solutions combinée" ont été testées : l'une dans le domaine fréquentiel (PM), l'autre dans le domaine temporel (BACCPM). Bien que les deux formulations conduisent à des performances quasi identiques, les filtres résultants de la méthode BACCPM présentent un « pre-ringing » plus faible que celui obtenu avec la méthode PM. Cependant, le temps de calcul des filtres obtenus dans le domaine temporel est 1000 fois plus important que celui des filtres obtenus dans le domaine fréquentiel.

L'approche déterministe habituellement employée pour la mise en œuvre des PSZ suppose que l'environnement acoustique est invariant dans le temps (que les réponses impulsionnelles utilisées pour le calcul sont valables quelque soient les conditions). Cependant, l'habitacle d'une voiture est un environnement acoustiquement variable où de nombreuses modifications peuvent se produire et modifier la réponse acoustique du système.

C'est pourquoi la sensibilité du système aux modifications typiques qui peuvent se produire dans un habitacle de voiture a été évaluée expérimentalement. À l'aide du prototype étudié dans la première partie, trois modifications ont été choisies comme étant les plus susceptibles de se produire dans un habitacle de voiture : la position des sièges, la hauteur de l'appui-tête et la présence de passagers sur les sièges arrière.

Il apparaît que l'environnement acoustique joue un rôle important sur les performances du système. C'est pourquoi le système a été testé en considérant chaque modification comme un nouvel environnement, c'est-à-dire en appliquant à nouveau toutes les étapes déterministes développées dans la première partie (appelées dans ce mémoire "approche avec ré-étalonnage"). Dans ces conditions, le système PSZ proposé est capable de répondre aux contraintes d'origine, quelle que soit la modification acoustique introduite (indépendamment du positionnement du siège, du positionnement de l'appui-tête et de la présence de passagers aux places arrière).

Néanmoins, afin d'obtenir les performances souhaitées, un ensemble de filtres spécifiques doit être utilisé pour chaque nouvel environnement acoustique. Il a été observé que si un filtre identique est appliqué dans toutes les conditions (ce qu'on appelle l'approche fixe), le contraste acoustique peut diminuer de 15 dB, et dans certains cas être inférieur au contraste naturel obtenu sans traitement.

Le défi consistait donc à trouver le moyen d'obtenir la combinaison optimale de filtres lorsqu'il y a une modification de l'habitacle de la voiture. Pour cela différentes stratégies sont proposées dans cette thèse et sont présentées ci-dessous avec leurs avantages et inconvénients. Elles constituent deux familles : celles qui découlent de l'approche déterministe et celles qui mettent en œuvre une approche adaptative (voir Fig. R.3).

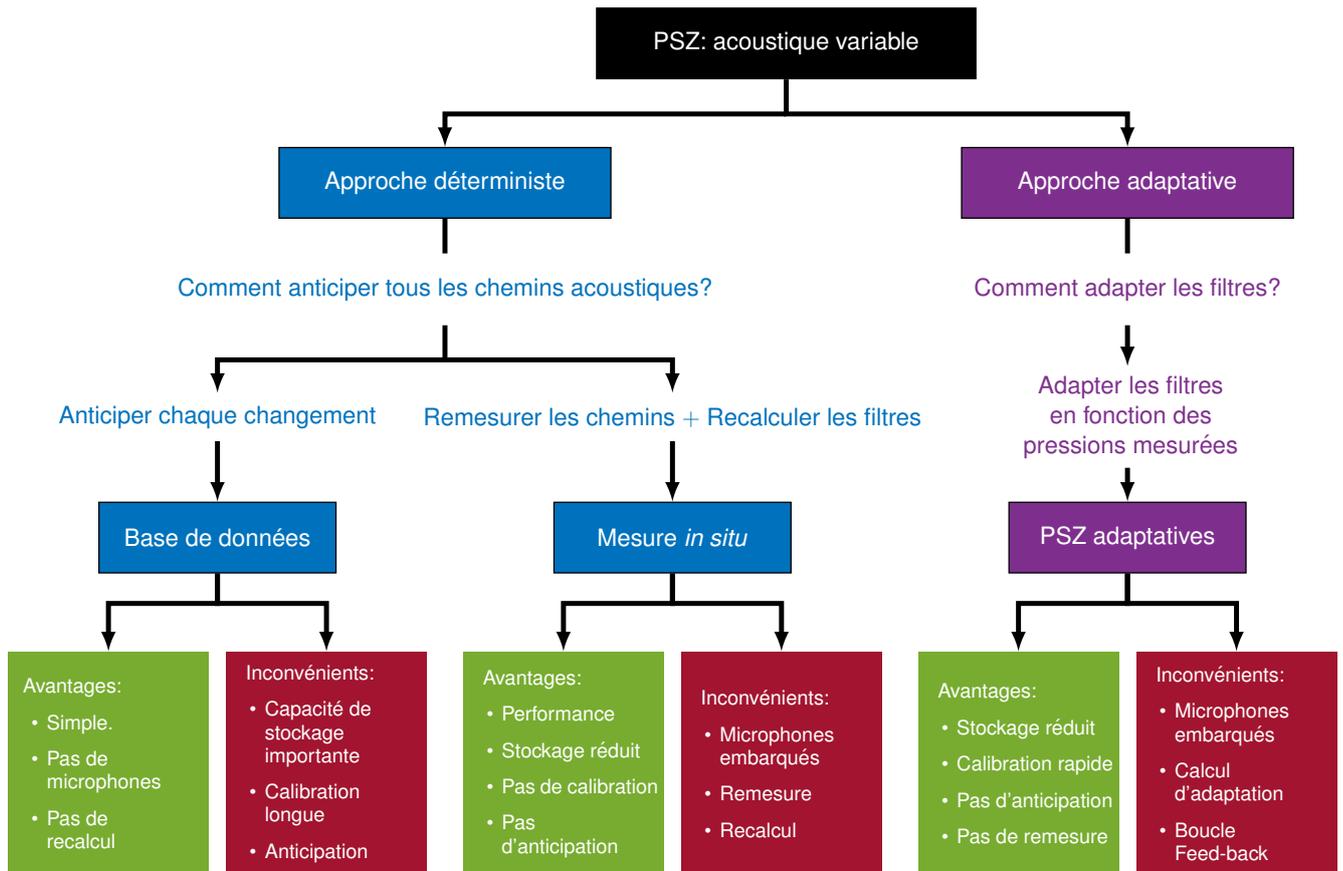


Figure R.3: Approches utilisées dans cette thèse pour prendre en compte la variabilité acoustique de l'habitacle d'une voiture.

### Approche déterministe

Deux solutions directement issues de l'approche déterministe sont envisagées : la solution utilisant une base de données et la solution utilisant une mesure in situ. Pour chacune de ces solutions, l'enjeu est d'obtenir une caractérisation précise de l'environnement acoustique pour calculer les filtres optimaux et ainsi obtenir les performances optimales du système.

La première option envisagée est celle dite de la base de données. Cette solution consiste à caractériser expérimentalement la réponse du système audio dans l'habitacle de voiture pour toutes les modifications acoustiques possibles et à appliquer les différentes étapes de traitement pour chaque modification. Tous les filtres optimaux ainsi obtenus sont stockés et choisis en fonction d'autres capteurs (caméra, position électrique du siège, etc.). En supposant que l'on puisse caractériser chaque configuration en fin de chaîne de production, il est nécessaire de disposer d'une unité de stockage de toutes les combinaisons de filtres dans une voiture.

La deuxième option envisagée est la solution de la mesure in situ. Cette solution repose sur le principe suivant : lorsqu'une perte de performance est détectée, le système audio est stoppé et les étapes de l'approche déterministe sont mises en œuvre en cours d'utilisation du véhicule. L'un des principaux enjeux de cette solution est que, pour mesurer les réponses impulsionnelles, il est nécessaire d'installer des microphones dans l'habitacle de la voiture. Comme il n'est pas possible d'utiliser la configuration d'origine (36 microphones situés près des têtes), une étude a été réalisée afin de réduire le nombre de microphones. Les résultats montrent qu'il est possible de réduire le nombre de microphones à deux par oreille (8 microphones au total) sans détériorer (ou presque) les performances.

Par exemple, dans la figure suivante, on peut observer les résultats de contraste acoustique ( $AC$ ) et d'erreur ( $E_{rr}$ ) obtenus en utilisant huit microphones, pour cinq positions de siège différentes lorsque les

deux sièges sont déplacés symétriquement (+12 cm et +6 cm vers l'avant de la voiture, -6 cm et -12 cm vers l'arrière de la voiture).

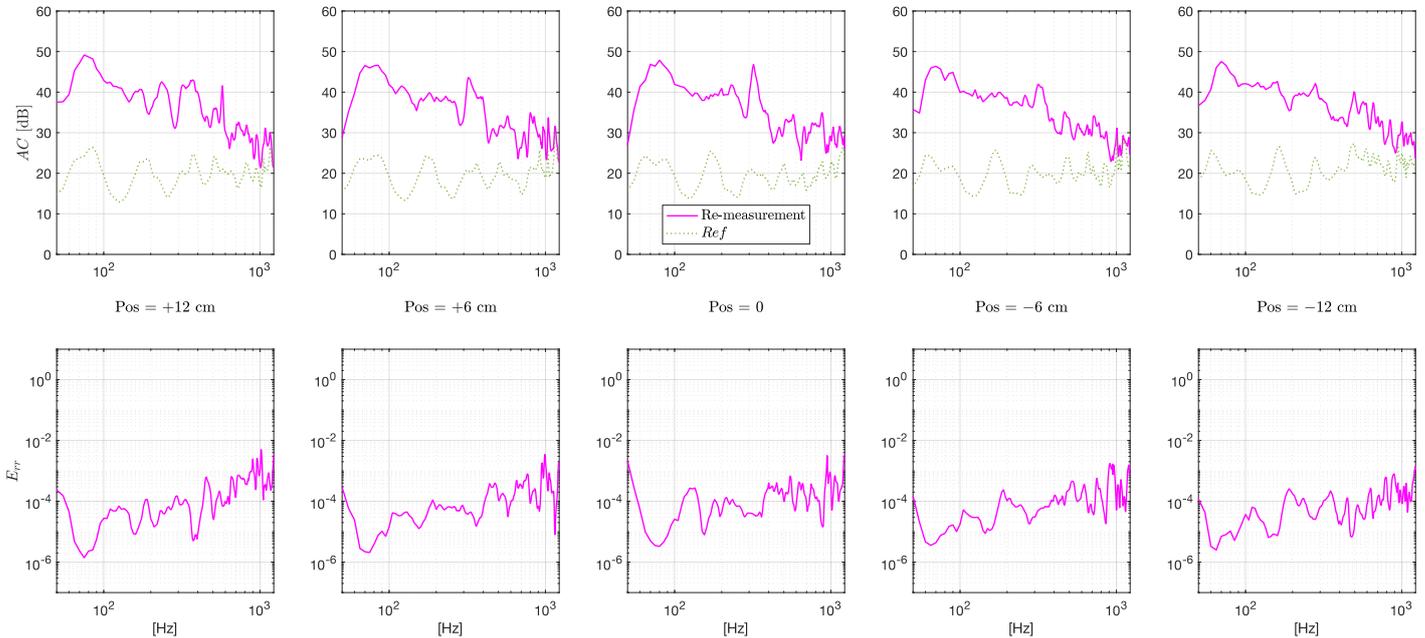


Figure R.4: Contraste Acoustique ( $AC$ ) et erreur ( $E_{rr}$ ) obtenus pour cinq positions de sièges symétriques en utilisant deux microphones par oreille (au total huit microphones). Chaque colonne correspond à une position de siège.

Afin de rendre ces solutions plus adaptées à l'application automobile, la méthode « Forced Pressure Matching » (FPM) a été développée. Cette méthode suppose que certains filtres sont imposés et permet de repenser la méthode du « pressure matching ». L'utilisation de la méthode FPM permet de réduire de 25% la quantité de stockage de la base de données et le temps de calcul d'environ 21%, ceci sans compromettre les performances.

### Approche adaptative

Une alternative à l'approche déterministe consiste à utiliser des algorithmes de filtrage adaptatifs pour maîtriser la variabilité acoustique de l'habitacle de la voiture. Le principe de ces méthodes est d'adapter itérativement les filtres en fonction de la pression résultante dans les zones. Dans cette thèse, l'algorithme FxLMS a été étudié en raison de sa simplicité et de son faible coût de calcul.

La caractéristique principale de l'implémentation de FxLMS est que, pour calculer les filtres, il ne faut qu'une connaissance imparfaite de la réponse de l'habitacle (un modèle approché du système suffit). Par conséquent, il a été supposé que les réponses impulsionnelles obtenues dans une configuration acoustique particulière (appelée configuration d'étalonnage) permettent de calculer les filtres pour n'importe quelle configuration.

L'une des limites de l'algorithme FxLMS est qu'il présente une convergence dépendant de la fréquence, ce qui compromet la performance du système dans les basses fréquences. Par conséquent, dans cette thèse, il est proposé de surmonter cette limitation en utilisant comme point de départ de l'algorithme adaptatif (comme coefficients initiaux), les filtres optimaux calculés dans la configuration d'étalonnage à l'aide de l'approche déterministe.

Les résultats obtenus montrent que l'algorithme FxLMS est capable d'obtenir de bonnes performances dans toutes les configurations étudiées. L'utilisation d'un seul modèle du système (mesuré pour une seule position) permet d'atteindre des niveaux de contraste acoustique ( $AC$ ) et d'erreur ( $E_{rr}$ ) similaires à ceux

issus de l'approche avec « ré-étalonnage » pour de petites modifications. Dans des positions plus extrêmes, le système reste performant, mais il n'atteint pas les performances de l'approche avec « ré-étalonnage ».

Néanmoins, comme pour les méthodes avec « ré-étalonnage », l'approche adaptative nécessite l'installation de microphones dans l'habitacle afin de recalculer les filtres. Les résultats montrent que le système adaptatif permet de conserver ses performances en utilisant seulement deux microphones par oreille (un total de huit microphones).

Enfin, il a également été démontré qu'il est possible d'étendre les principes du FPM à l'algorithme FxLMS, non seulement pour réduire le nombre de coefficients à adapter, mais aussi pour certaines configurations de microphones qui permettent de diminuer l'erreur. Cette méthode est appelée l'algorithme FxLMS forcé. Par exemple, la figure R.5 montre le contraste acoustique ( $AC$ ) et l'erreur ( $E_{rr}$ ) pour cinq positions de siège différentes (les deux sièges sont déplacés symétriquement tous les 6 cm) en utilisant l'algorithme FxLMS forcé après 10 secondes de calcul en utilisant deux microphones par oreille.

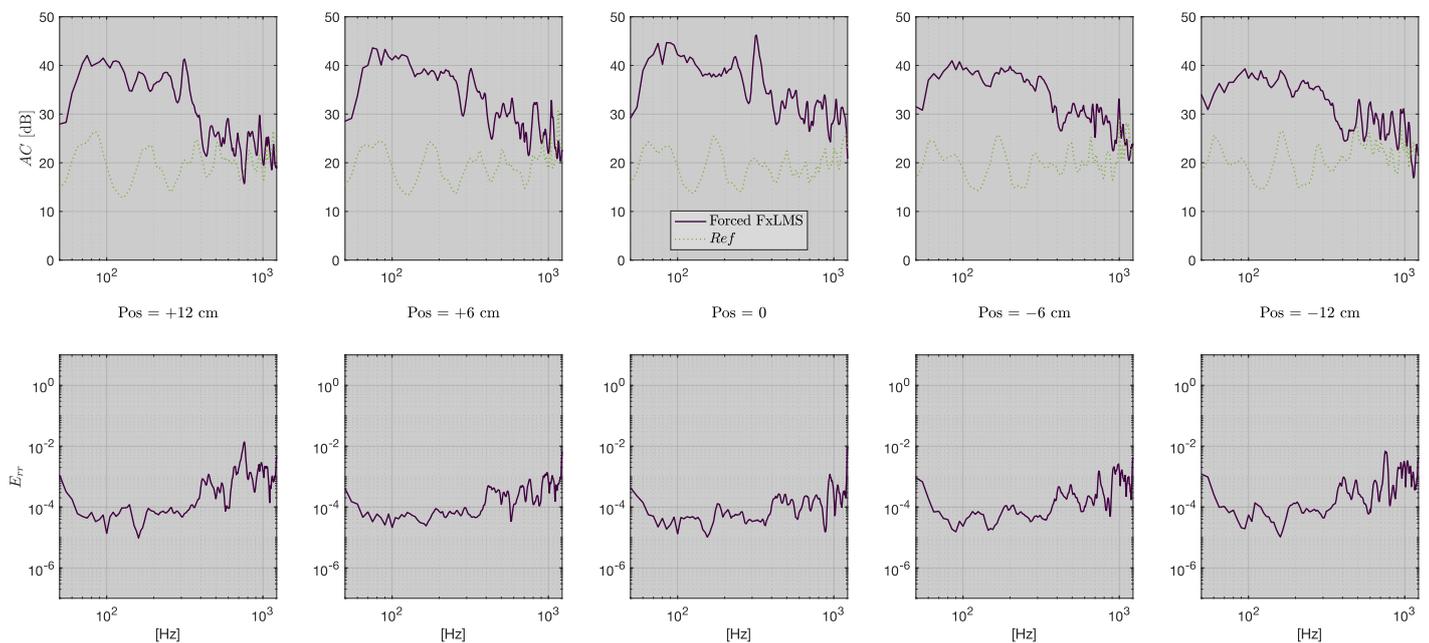


Figure R.5: Contraste Acoustique ( $AC$ ) et erreur ( $E_{rr}$ ) obtenus grâce à l'algorithme FxLMS forcé avec huit microphones pour cinq positions de sièges symétriques. Chaque colonne correspond à une position de siège..

En conclusion, les exigences initiales établies au début de la thèse ont été respectées. Un système fonctionnant dans un habitacle automobile a été étudié et différentes options ont été proposées pour maintenir les performances du système dans le cas de modifications usuelles de l'habitacle.

**Titre : Personalised low-frequency sound reproduction in a car cabin**

**Mots clés :** Personal sound zones, Car cabin, Constrained optimisation, Adaptive filters.

**Résumé :** The generation of personal sound zones has been a topic of great interest for more than twenty years, with the automotive industry as one of its main fields of application. The goal of this thesis is to study the feasibility of a two-zone system, below 1 kHz, capable of maintaining its performance to possible changes that might happen in a car cabin.

The first part of this thesis focuses on designing a system capable of achieving an acoustic contrast (the ratio between the zones acoustic potential energy) around 30 dB, whilst maintaining a low normalised reproduction error in the listening zone.

An eight-loudspeaker system was designed consisting of a four-loudspeaker custom headrest on each seat. Furthermore, the prototype was built and tested experimentally in three different acoustic environments: anechoic conditions, anechoic conditions with reflective panels and a car cabin. Using the pressure matching algorithm, the prototype achieved a mean acoustic contrast of 34.7 dB and a mean reproduction error of  $1.7 \times 10^{-4}$  over the studied bandwidth, in a car cabin.

The second part of this dissertation tackles the inherent acoustic variability of a car cabin. Using the aforementioned prototype, three different modifications were chosen to be studied experimentally: position of the seats, elevation of the headrest and the addition of passengers in the back seats. The prototype was capable of maintaining its performance regardless of the acoustic modification introduced. However, a dedicated set of filters had to be used for each “new-acoustic conditions”.

Two different alternatives to obtain these dedicated filters were reviewed, to anticipate every acoustic modification (data-base solution) or to re-characterise the car cabin with a few microphones and recalculate the filters (re-measurement solution). In order to make these solutions more attractive for the car application, the forced pressure matching (FPM) method was introduced. This method is a rethinking of the pressure matching method when some filter responses are imposed in the calculation. The use of the FPM for the application in hand, reduces the amount of storage needed by the database-solution by 25% without compromising the performance. Additionally, the FPM reduces the calculation time of the filters by  $\approx 21\%$  without compromising the performance.

Lastly, as a third alternative, the use of adaptive filter algorithms is proposed (specifically the widely-studied filtered-x least mean squares). The key feature of its implementation is that, in order to calculate the filters, it only requires an imperfect knowledge of the car cabin (plant model).

The method proposed in this dissertation assumes that it is possible to characterise the car cabin only once, for one position. Moreover, this information is used to define one plant model and to calculate the algorithm's initial coefficients (using the pressure matching algorithm). The method showed to be capable of improving the performance of the system regardless of the modification. Lastly, a thorough study on the reduction of microphones installed in the car is performed, resulting in a successful implementation of the system with only eight error microphones.