Alma Mater Studiorum - Università di Bologna

Dottorato di Ricerca in

Automotive per una Mobilità Intelligente

Ciclo 34

Settore Concorsuale: 09/F2

Settore Scientifico Disciplinare: ING-INF/03 Telecomunicazioni

Active Control of the Acoustic Field in a Vehicle Cabin

Presentata da: Alessandro Opinto

Coordinatore Dottorato
Prof. Nicolò Cavina

Supervisore
Prof. Angelo Farina

Co-supervisori
Prof. Stefano Selleri
Prof. Riccardo Raheli

Esame finale anno 2022

Table of Contents

List of Acronyms

AD	Automatic Differentiation
ANC	Active Noise Control
BBPF	Broad Band-Pass Filter
BPF	Band-Pass Filter
CAN	Controller Area Network
CICO	Combined Input - Combined Output
CNN	Convolutional Neural Network
DC	Direct Current
DNN	Deep Neural Network
DSP	Digital Signal Processor
EOC	Engine Order Cancellation
FB	FeedBack
FCN	Fully Convolutional Network
FE	Filtered Error
FF	FeedForward
FIR	Finite Impulse Response
FxLMS	Filtered-x Least Mean Square

GPU	Graphical Processing Unit
IMC	Internal Model Control
KPI	Key Parameter Indicator
LMS	Least Mean Square
LPF	Low-Pass Filter
MIMO	Multiple Input - Multiple Output
MISO	Multiple Input - Single Output
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
NBPF	Narrow Band-Pass Filter
NN	Neural Network
NR	Noise Reduction
OF	Observation Filter
RBW	Resolution BandWidth
ReLU	Rectified Linear Unit
RLS	Recursive Least Squares
RMT	Remote Microphone Technique
RNC	Road Noise Cancellation
rpm	Revolution Per Minute
SIMO	Single Input - Multiple Output
SISO	Single Input - Single Output
SPL	Sound Pressure Level
SVD	Singular Value Decomposition

TCN Temporal Convolutional Network

VMT Virtual Microphone Technique

WGN White Gaussian Noise

List of Symbols

English Letters

SYMBOL	DESCRIPTION
d, D	monitoring microphone disturbing signal, transform of d
e, E	monitoring microphone error signal, transform of e
f	frequency
f_c	central frequency of octave band filter bank
j, J	reference signal index, number of reference signals
k, K	loudspeaker index, number of loudspeakers
m, M	monitoring microphone index, number of monitoring microphones
n	discrete-time epoch
o, O	observation filter impulse response, transform of o
p, P	primary path impulse response at monitoring microphone, transform of p
s, S	secondary path impulse response at monitoring microphone, transform of s
T	input/output transfer function
u	unknown audio wave
v, V	virtual microphone index, number of virtual microphones
w, W	control filter impulse response, transform of w
x, X	reference signal, transform of x
y, Y	anti-noise signal at monitoring microphone, transform of y

Greek Letters

SYMBOL	DESCRIPTION
α	positive constant to prevent division by zero for LMS algorithm
Γ	effective cancellation
δ, Δ	virtual microphone disturbing signal, transform of d
ε	error signal between virtual microphone signal and its retrieved version
η	theoretical cancellation
ϑ	dilation factor for TCN
λ	leakage-factor
μ	step-size parameter
ξ, Ξ	virtual error microphone signal, transform of ξ
π, Π	primary path impulse response at virtual microphone, transform of π
ρ	regularization factor for MMSE algorithm
ϱ	neuron for NN
σ, Σ	secondary path impulse response at virtual microphone, transform of σ
Υ	MSE evaluated on error signal $\varepsilon[n]$
χ	cost function
ψ, Ψ	anti-noise signal at virtual microphone, transform of ψ
ω	angular frequency

Foreword

This thesis describes the research activity carried on during the PhD program in "*Automotive Engineering for Intelligent Mobility*" and is focused on the analysis and design of Active Noise Control (ANC) systems for automotive applications. This work was supported by ASK Industries S.p.A., Reggio Emilia, Italy.

Abstract

In this thesis, a thorough investigation on acoustic noise control systems for realistic automotive scenarios is presented. The thesis is organized in two parts dealing with the main topics treated: Active Noise Control (ANC) systems and Virtual Microphone Technique (VMT), respectively.

The technology of ANC allows to increase the driver's/passenger's comfort and safety exploiting the principle of mitigating the disturbing acoustic noise by the superposition of a secondary sound wave of equal amplitude but opposite phase. Performance analyses of both FeedForwrd (FF) and FeedBack (FB) ANC systems, in experimental scenarios, are presented. Since, environmental vibration noises within a car cabin are time-varying, most of the ANC solutions are adaptive. However, in this work, an effective fixed FB ANC system is proposed. Various ANC schemes are considered and compared with each other. In order to find the best possible ANC configuration which optimizes the performance in terms of disturbing noise attenuation, a thorough research of Key Parameter Indicator (KPI), system parameters and experimental setups design, is carried out.

In the second part of this thesis, VMT, based on the estimation of specific acoustic channels, is investigated with the aim of generating a quiet acoustic zone around a confined area, e.g., the driver's ears. Performance analysis and comparison of various estimation approaches is presented. Several measurement campaigns were performed in order to acquire a sufficient duration and number of microphone signals in a significant variety of driving scenarios and employed cars. To do this, different experimental setups were designed and their performance compared. Design guidelines are given to obtain good trade-off between accuracy performance and equipment costs. Finally, a preliminary analysis with an innovative approach based on Neural Networks (NNs) to improve the current state of the art in microphone virtualization is proposed.

Part I

Active Noise Control Systems

Introduction

In the last decades, due to the technological improvements and the consequent increase in the number of equipments, such as engines, blowers, fans and compressors, noise control systems have become a hot research topic in the automotive industry. For the medium-and high-frequency disturbing signals, the problem of noise mitigation within the car cabin is normally approached via "passive" mitigation methods, e.g., by introducing acoustic absorption materials. However, these solutions may be ineffective for the low-frequency disturbing audio wave [1, 2]. To improve comfort, Active Noise Control (ANC) systems are being currently developed. The idea behind ANC is to cancel the unwanted (primary) noise source by emitting an anti-noise (secondary) of equal amplitude and opposite phase such that the residual noise is reduced as much as possible by the superposition principle [3–5]. Depending on the presence of a reference signal, we can distinguish between a FeedForward (FF) and a FeedBack (FB) ANC system. The FF ANC system operates in the presence of reference control signals, typically obtained by accelerometers placed in crucial positions of the car structure, loudspeakers emitting the anti-noise, and microphones, referred to as error microphones, within the cabin, employed to constantly monitor the residual noise perceived by the driver and the passengers. On the other hand, the FB ANC systems, that are based on the Internal Model Control (IMC) architecture [6], in the absence of reference signals, try to cancel the overall noise in the cabin by using the error microphone signals as references.

Since acoustic noise sources and the environment are time-varying, ANC must be adaptive. The most common solution is to employ a transversal filter using the Least Mean Square (LMS) algorithm [1]. The system will thus adjust the filter coefficients to attempt proper setting of the physical filters between the loudspeakers and the error microphones. Ideally, at the error microphones we therefore obtain residual errors that

approach zero.

In this context, the aim of reference signals in the FF ANC systems is to pick up the unwanted noise component. Therefore, the use of reference signals can be seen as an aid to predict what is listened by the error microphones [1]. This means that a necessary condition to have noise cancellation is that the reference signals have to be correlated with the unwanted noise component. When the number of employed reference signals is larger than one, the ANC system is referred to as multi-reference. ANC systems can be classified on the basis of the number of used loudspeakers and microphones as Single Input - Single Output (SISO), Multiple Input - Single Output (MISO) or Multiple Input - Multiple Output (MIMO). More precisely, in a SISO system we have a single loudspeaker and one microphone. In a MISO system, we have multiple loudspeakers and one single microphone. By employing multiple loudspeakers and more than one microphone, we finally obtain a MIMO system.

The goal of this Part I is the design and performance analysis of digital signal processing algorithms suited for the acoustic control in the cabin of a car. In the first part a thorough investigation of the system performance of FF ANC in automotive environments in the case of SISO, MISO, and MIMO scenarios is given. In particular, multiple reference signals are employed and investigated in order to find a specific reference or a grouping of references which maximize the performance. System performance is assessed in terms of Sound Pressure Level (SPL), i.e., the sound level perceived by the driver and/or passengers. Moreover, in order to investigate how the noise cancellation performs in time, the spectrogram is considered. Since in the car environment the noise sources are usually dominated by low frequencies, we expect that the noise cancellation will occur in the range 20–500 Hz.

Similarly to the FF case, in the FB approach adaptive LMS-based solutions have been analyzed. Systems were evaluated in terms SPL of the perceived audio, both its spectrum and time-domain sliding window values. Moreover, the problem of the design of a fixed FB controller is tackled. Unlike adaptive solutions, such controllers do not track the audio dynamics, but try to perform "static" cancellation. We hope that FB systems may complement FF ones for narrow band noise components not present in the reference signals. We expect that, the FB ANC system, due to the lack of reference signals may allow to cancel additional narrow-band components of the incoming noise.

Part I is structured as follows. The ANC problem is stated in Chapter 1 in terms of

the optimal Wiener solution and the adopted LMS algorithm. Chapter 2 summarizes the simulation setup and specifics implemented in our simulations. In Chapter 3, the FF ANC system is analyzed for the SISO, MISO and MIMO schemes. The considered solutions for FB ANC system, either adaptive or fixed, are presented in Chapter 4. In Chapter 5, considered input signals for the analyzed FB ANC schemes are presented. Simulation results are presented in Chapter 6 and organized according to the considered secondary path. In particular, three types of secondary paths were taken into consideration: modeling pure delays, representing a realistic cabin of a car and describing the behavior of the headrest of a chair. The obtained results are discussed in Sections 6.2, 6.3 and 6.4 respectively. Conclusions and future work are finally discussed.

State of the Art

The history of ANC started in 1936 thanks to the first patent of noise control system invented by Paul Lueg, a doctor of philosophy and medicine, who worked also as a physicist [7]. The idea of Paul Lueg was to mitigate sinusoidal tones in ducts by phase-advancing the wave and canceling arbitrary sounds in the region around a loudspeaker by inverting the polarity. In the early 1950s, the electrical engineer Lawrence Jerome Fogel concentrated his efforts to develop the first noise cancellation headphone in the aviation field [8]. This system was originally created to protect the pilot's hearing system from the disturbing sound of the plane engines. Between 1956 and 1957, Willard Meeker, leader of the Air-Force Research Laboratory, known for the introduction of earmuffs and earplug in the pilot's helmet as a passive noise mitigation, in order to improve pilot's comfort, decided to develop a new headset in which also the ANC system was implemented [9].

In the mid '80s, a significant technology improvement in the ANC field was achieved, thanks to the contribution of Dr. Amar Bose, the founder of Bose Corporation. In fact, in 1989, the first headphones with ANC system by Bose, the so-called "Bose Aviation Headset X" were made available commercially [10]. Similarly, one of the main Bose competitor, the German production company Sennheiser developed the "LHM 45 NoiseGard", the world's first pilot's headset with ANC to receive FAA-TSO certification (Technical Standard Order of the Federal Aviation Administration) for Lufthansa airlines [11]. From here on, several companies started to develop and sell headphones with ANC system.

As described, first developments on the active noise canceling field were pursued in both the army and civil aviation fields by employing pilot's helmets. We had to wait until the early '90s to have the first ANC system in the automotive scenario. In fact, in 1992, the English car manufacturer Lotus developed the first Road Noise Cancellation

(RNC) system featured in the Nissan Bluebird [12]. Several years later, in 2009, Lotus started a collaboration with Harman Cardon in order to develop new RNC and Engine Order Cancellation (EOC) systems [13]. Nowadays, in the automotive field, the main goal of commercial ANC solutions is to cancel the most important engine harmonics.

Within the car cabin, different kinds of disturbing audio sources are present, such as the rolling of the wheel on the asphalt, wind-noise and engine noise [14]. Each of these audio waves has peculiar intrinsic acoustic characteristics that depend on several physical and environmental conditions, such as the car speed, the presence or not of passengers, the temperature and so on. The wind-noise component, e.g., becomes predominant above 100 km/h and it is generally broadband (50–500 Hz) [1]. Conversely, the sound produced by the internal combustion engine is usually tonal and the fundamental frequency f_0 can be evaluated as follows

$$f_0 = \frac{\text{rotation speed} \cdot \#\text{ of cylinders}}{60 \cdot \#\text{ of cycle}} \quad [\text{Hz}]$$

where the rotation speed is in Revolution Per Minute (rpm). Hence, for instance, under the assumption of a 4 stroke engine with 2 cylinders, which fires each cylinder only once every two crank revolutions (i.e., # of cycle = 2), and 6000 rpm, the fundamental frequency is $f_0 = 100$ Hz [15]. Based on the characteristics of the unwanted disturbing audio field, it is therefore possible to distinguish between broad and narrow-band ANC systems. It is worth noting that, narrow-band disturbing noise, since their periodicity is usually preserved, to deal with narrow-band noises is easier than broad-band ones. On the other hand, the assumption of periodic noise in the automotive environment is not typically satisfied. In fact, most of the times, the car cabin is affected by time-varying noises. For this reason, using adaptive ANC strategies, e.g., the Filtered-x Least Mean Square (FxLMS) algorithm, is often preferable with respect to fixed ones [3, 16].

With the aim of aiding the ANC system to mitigate the disturbing signals, sensors, e.g., microphones or accelerometers, can be installed in strategic positions of the car structure in order to acquire reference signals. As alternative, since the main vibration contribution propagating within the car cabin is due to the speed of the car, reference signals can be directly obtained by the tachometer, the Controller Area Network (CAN) bus or the ignition circuit of the car. When this strategy is considered, we talk about FF ANC system.

When reference signals are not available, the knowledge of the error microphone signals can be exploited to synthesize them. Since the error microphone signals have to be fed back to the Digital Signal Processor (DSP) for signal synthesis purposes, this scheme is called FB. The adaptive FB ANC systems are based on the IMC architecture and when the reference reconstruction is performed they are equivalent to the FF ones [6].

Based on the nature and the number of acoustic fields to be controlled by active mitigation systems, different ANC schemes may be employed. A SISO system may be sufficient to reduce periodic acoustic waves. On the other hand, MIMO systems may be introduced in order to weaken different disturbing signals in more than one position. Multiple loudspeakers can be employed to reduce disturbing waves in a single position, yielding a MISO ANC system. Finally, the Single Input - Multiple Output (SIMO) ANC system, not widely use due to physical limitations, exploits one loudspeaker to create several zones of silence for multiple error microphones.

Chapter 1

Adaptive Filter Estimation

The goal of this chapter is to illustrate the most common techniques for adaptive filter parameter estimation, since they represent a fundamental pillar in ANC systems [17].

The rest of this chapter is structured as follows. The main concepts on the Wiener filter are discussed in Section 1.1. An adaptive Wiener filter, based on an iterative approach named steepest-descent algorithm is presented in Section 1.2. Finally, in Section 1.3 the LMS method is derived.

1.1 Wiener Filter

Consider the block diagram in Fig. 1.1, where the causal time-invariant filter $W(z)$ such that the *error* signal $e[n]$ is minimized with respect to the output of a given filter $P(z)$ is desired. This means that $W(z)$ can be interpreted as an estimate of $P(z)$.

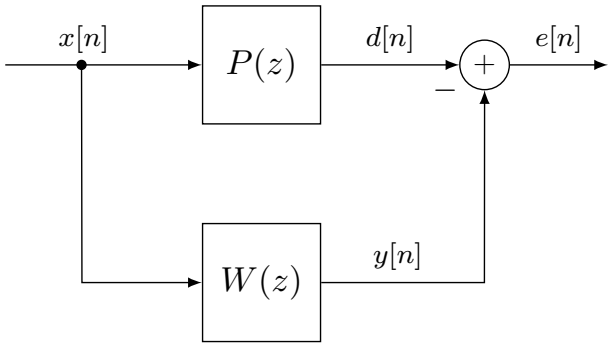


Figure 1.1: System model: filtering estimation scheme.

Assume a real filter input $x[n] \in \mathbb{R}$ and a low-pass Finite Impulse Response (FIR) filter $P(z)$. The error signal is defined as the difference between the so-called desired output $d[n]$ and the filter output $y[n]$, i.e.[1]

$$e[n] = d[n] - y[n] \ . \tag{1.1}$$

The Wiener filter applies the Minimum Mean Square Error (MMSE) criterion so that the error signal is minimized in the mean square sense. The cost function to be minimized, denoted as χ, is known as *mean-squared error*, and is defined as

$$\chi = \mathbb{E}\left\{e^2[n]\right\} \tag{1.2}$$

where $\mathbb{E}\{\cdot\}$ denotes the statistical expectation.

Consider now a transversal filter as shown in Fig. 1.2 with N taps and weights $w_i,\ i = 0, 1, ..., N-1$. Let us define the following N-length column vectors[2]: the input vector

$$\mathbf{x}[n] = [x[n], x[n-1], ..., x[n-N+1]]^\top$$

and the weight vector

$$\mathbf{w} = [w_0, w_1, ..., w_{N-1}]^\top$$

being $\top$ the transpose operator. The filter output $y[n]$ is

$$y[n] = \sum_{i=0}^{N-1} w_i x[n-i]$$

and can be also expressed in a more compact form as

$$y[n] = \mathbf{w}^\top \mathbf{x}[n] = \mathbf{x}^\top[n]\mathbf{w} \ . \tag{1.3}$$

[1] The error signal could be equivalently defined as $e[n] = y[n] - d[n]$, as also customary in the literature on adaptive filtering, provided the following derivation is consistently developed.

[2] In the remainder of this report, lowercase and uppercase bold letters denote vectors and matrices, respectively.

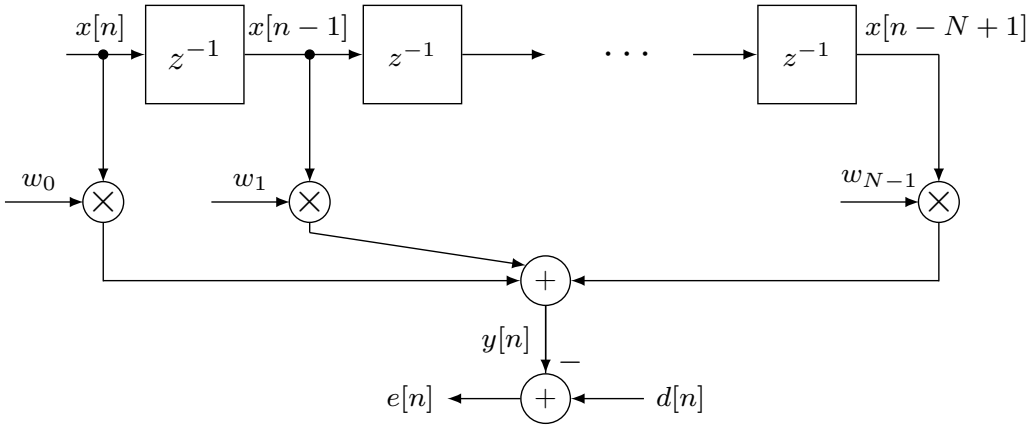


Figure 1.2: Transversal filter with N taps.

By substituting (1.3) into (1.1), one has

$$e[n] = d[n] - \mathbf{w}^\top \mathbf{x}[n] = d[n] - \mathbf{x}^\top[n]\mathbf{w} \ .$$

Using (1.2), the cost function χ can be written as

$$\begin{aligned}\chi &= \mathbb{E}\{e^2[n]\} \\ &= \mathbb{E}\{\big(d[n] - \mathbf{w}^\top \mathbf{x}[n]\big)\,\big(d[n] - \mathbf{x}^\top[n]\mathbf{w}\big)\} \\ &= \mathbb{E}\{d^2[n]\} - \mathbf{w}^\top \mathbb{E}\{\mathbf{x}[n]d[n]\} - \mathbb{E}\{d[n]\mathbf{x}^\top[n]\}\mathbf{w} + \mathbf{w}^\top \mathbb{E}\{\mathbf{x}[n]\mathbf{x}^\top[n]\}\mathbf{w} \ . \end{aligned} \quad (1.4)$$

The following cross-correlation vector can be defined:

$$\mathbf{p} = \mathbb{E}\{\mathbf{x}[n]d[n]\} = [p_0, p_1, ..., p_{N-1}]^\top$$

where $p_i = \mathbb{E}\{x[n-i]d[n]\}$ for $i = 0, 1, ..., N-1$. The following auto-correlation matrix can also be defined:

$$\mathbf{R} = \mathbb{E}\{\mathbf{x}[n]\mathbf{x}^\top[n]\} = \begin{bmatrix} r_{0,0} & r_{0,1} & \cdots & r_{0,N-1} \\ r_{1,0} & r_{1,1} & \cdots & r_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N-1,0} & r_{N-1,1} & \cdots & r_{N-1,N-1} \end{bmatrix}$$

where $r_{i,j} = \mathbb{E}\{x[n-i]x[n-j]\}$ for $i, j = 0, 1, ..., N-1$. $\mathbf{R}$ is obviously symmetric, as

$\mathbf{R} = \mathbf{R}^\top$. Therefore, (1.4) can be rewritten in a more compact form as

$$\chi = \mathbb{E}\{d^2[n]\} - 2\mathbf{w}^\top\mathbf{p} + \mathbf{w}^\top\mathbf{R}\mathbf{w} \tag{1.5}$$

in which the fact that $\mathbf{w}^\top\mathbf{p} = \mathbf{p}^\top\mathbf{w}$ was exploited. In order to minimize (1.5), we can setup the following minimization problem:

$$\frac{\partial\chi}{\partial w_i} = 0 \qquad i = 0, 1, ..., N-1$$

or, equivalently,

$$\nabla\chi = \mathbf{0} \tag{1.6}$$

being $\mathbf{0}$ the all-zero column vector of size N and ∇ the gradient defined as the column operator

$$\nabla \triangleq \left[\frac{\partial}{\partial w_0}, \frac{\partial}{\partial w_1}, ..., \frac{\partial}{\partial w_{N-1}}\right]^\top .$$

By plugging (1.5) into (1.6), one has

$$\nabla\chi = 2\mathbf{R}\mathbf{w} - 2\mathbf{p} = \mathbf{0} \tag{1.7}$$

which leads to

$$\mathbf{R}\mathbf{w} = \mathbf{p} \ . \tag{1.8}$$

Equation (1.8) is referred to as *Wiener-Hopf* and its solution, known as *Wiener-Hop solution*, is

$$\mathbf{w}_0 = \mathbf{R}^{-1}\mathbf{p} \tag{1.9}$$

where $\mathbf{w}_0$ denotes the optimal tap weight vector. Equation (1.9) represents the filter tap weight vector which minimizes the mean square error.

Note that the auto-correlation matrix $\mathbf{R}$ has to be invertible to make (1.9) valid.

1.2 Steepest Descent Method

In the previous section, the optimum tap weights of a transversal Wiener filter was found by directly solving the Wiener-Hopf equation (1.8). An alternative method

is represented by an iterative search, through the well-known *steepest descent algorithm* [3, 18].

The steepest descent algorithm is based on the gradient iterative method and can be implemented by following this procedure:

1. start with an initial guess of the parameters to be optimized

2. find the gradient of the cost function with respect to these parameters at the present point

3. update the parameters by taking a step in the direction of the negative gradient

4. repeat steps 2. and 3. until no further significant changes are observed in the parameters.

To implement this method, consider again Fig. 1.2 and (1.7). Following the above procedure, the recursive equation to update the filter tap weight vector at the k-th iteration ($k = 1, 2, ...$) is

$$\mathbf{w}[k+1] = \mathbf{w}[k] - \frac{1}{2}\mu\nabla_k\chi \tag{1.10}$$

where μ is a positive scalar parameter known as *step-size* which controls the convergence speed of the algorithm, ∇_k denotes the gradient at time step k, and $\mathbf{w}[k] = [w_0[k], w_1[k], ..., w_{N-1}[k]]^\top$ denotes the time-varying tap weight vector. Substituting (1.7) into (1.10) one obtains

$$\begin{aligned}\mathbf{w}[k+1] &= \mathbf{w}[k] - \frac{1}{2}\mu\nabla_k\chi \\ &= \mathbf{w}[k] - \mu\left(\mathbf{R}\mathbf{w}[k] - \mathbf{p}\right) \\ &= \left(\mathbf{I} - \mu\mathbf{R}\right)\mathbf{w}[k] - \mu\mathbf{p}\end{aligned} \tag{1.11}$$

where $\mathbf{I}$ is the identity matrix of size $N \times N$. Equation (1.11) represents the filter tap-weight update equation. It can be proved that convergence is guaranteed if $0 < \mu < 2/\lambda_{max}$ where, $\lambda_{\max} \triangleq \max_{i=0,1,...,N-1} \lambda_i$, being λ_i the eigenvalues of the auto-correlation matrix $\mathbf{R}$ [17].

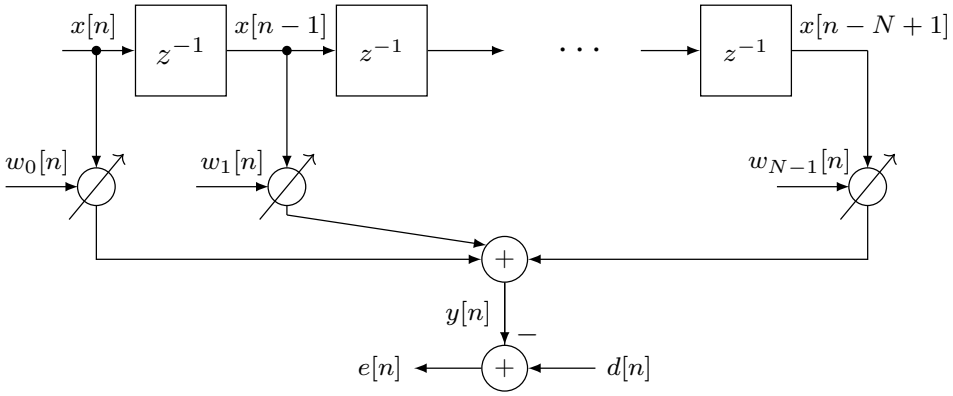


Figure 1.3: Transversal adaptive filter with time-varying weights.

1.3 Least Mean Square Algorithm

The main drawback of previous methods is the time-invariant assumption on the filter. Unfortunately, in general, the tap weights are time-varying. For this reason, another method, which takes into account the time-varying nature of a filter, is the LMS algorithm.

The LMS algorithm adapts the filter tap weights by continuous observation of its input $x[n]$ so that the mean square error $\mathbb{E}\{e^2[n]\}$ is minimized. From a practical point of view, LMS can be seen as a stochastic implementation of the steepest descent algorithm in a convex problem. In this section, the standard LMS algorithm for a transversal adaptive filter is derived. Moreover, two modified versions of the standard LMS, namely Normalized and Leaky LMS, are described.

Standard LMS

Let us consider the transversal adaptive filter in Fig. 1.3. The main difference between the Wiener solution and LMS is in the cost function. In fact, if in Section 1.1 the cost function is $\chi = \mathbb{E}\{e^2[n]\}$, it is here substituted by its instantaneous estimate

$$\hat{\chi}[n] = e^2[n] \ . \tag{1.12}$$

By adapting (1.10) to the cost function (1.12) the following update recursion is obtained:

$$\mathbf{w}[n+1] = \mathbf{w}[n] - \frac{1}{2}\mu\nabla_k e^2[n] \ . \tag{1.13}$$

The $i-$th element of the gradient of the squared error signal is evaluated as follows:

$$\begin{aligned}
\frac{\partial e^2[n]}{\partial w_i} &= 2e[n]\frac{\partial e[n]}{\partial w_i} \\
&= 2e[n]\frac{\partial}{\partial w_i}\left(d[n] - y[n]\right) \\
&= -2e[n]\frac{\partial y[n]}{\partial w_i} \\
&= -2e[n]\underbrace{\frac{\partial}{\partial w_i}\sum_{\ell=0}^{N-1} w_\ell x[n-\ell]}_{\neq 0 \text{ only when } \ell=i} \\
&= -2e[n]x[n-i] \qquad i = 0, 1, ..., N-1
\end{aligned}$$

in which the independence between w_i and $d[n]$ and the fact that the derivative of the sum in the $4-$th line is non-zero for $l = i$, only, was exploited. In a vectorial form, one can compactly write:

$$\nabla e^2[n] = -2e[n]\mathbf{x}[n] \ . \tag{1.14}$$

Now, by substituting (1.14) into (1.13) one finally obtains the filter tap weight update recursion of the LMS algorithm[3]

$$\boxed{\mathbf{w}[n+1] = \mathbf{w}[n] + \mu e[n]\mathbf{x}[n] \ .} \tag{1.15}$$

LMS Convergence

In Section 1.2, the convergence of the steepest descent method is mentioned, stating that it is guaranteed only for a specific range of the step-size parameter μ. In this subsection, the proof of the LMS convergence is derived.

Let us consider the update equation (1.15) and focus on the term $e[n]$. Assume that the optimum tap weight vector $\mathbf{w}_0$ is considered. By using (1.1), the error signal

[3]Had the definition of error signal in Footnote [1] been used, (1.13) would have obviously shown a minus sign in its right-hand side.

can be rearranged as follows:

$$\begin{aligned} e[n] &= d[n] - y[n] \\ &= d[n] - \mathbf{x}^\top[n]\mathbf{w}[n] \\ &= d[n] - \mathbf{x}^\top[n]\mathbf{w}[n] - \mathbf{x}^\top[n]\mathbf{w}_0 + \mathbf{x}^\top[n]\mathbf{w}_0 \\ &= d[n] - \mathbf{x}^\top[n]\mathbf{w}_0 - \mathbf{x}^\top[n]\left(\mathbf{w}[n] - \mathbf{w}_0\right) \\ &= e_0[n] - \mathbf{x}^\top[n]\mathbf{a}[n] \end{aligned} \tag{1.16}$$

where $e_0[n] \triangleq d[n] - \mathbf{x}^\top[n]\mathbf{w}_0$ is the estimation error obtained with the filter optimum tap weights and $\mathbf{a}[n] \triangleq \mathbf{w}[n] - \mathbf{w}_0$ is the weight error vector with respect to the optimum. Thus, by substituting (1.16) into (1.15) and subtracting $\mathbf{w}_0$ one has

$$\begin{aligned} \mathbf{a}[n+1] &= \mathbf{w}[n+1] - \mathbf{w}_0 \\ &= \mathbf{w}[n] - \mathbf{w}_0 + \mu\left(e_0[n] - \mathbf{x}^\top[n]\mathbf{a}[n]\right)\mathbf{x}[n] \\ &= \mathbf{a}[n] + \mu e_0[n]\mathbf{x}[n] - \mu\mathbf{x}[n]\mathbf{x}^\top[n]\mathbf{a}[n] \\ &= \left(\mathbf{I} - \mu\mathbf{x}[n]\mathbf{x}^\top[n]\right)\mathbf{a}[n] + \mu e_0[n]\mathbf{x}[n] \end{aligned}$$

which represents the tap weight update equation in terms of the vector $\mathbf{a}$. By taking the expectation of both sides, one obtain

$$\begin{aligned} \mathbb{E}\{\mathbf{a}[n+1]\} &= \mathbb{E}\{\left(\mathbf{I} - \mu\mathbf{x}[n]\mathbf{x}^\top[n]\right)\mathbf{a}[n]\} + \mu \underbrace{\mathbb{E}\{e_0[n]\mathbf{x}[n]\}}_{=0 \text{ by orth.}} \\ &= \mathbb{E}\{\mathbf{a}[n] - \mu\mathbf{x}[n]\mathbf{x}^\top[n]\mathbf{a}[n]\} \\ &= \mathbb{E}\{\mathbf{a}[n]\} - \mu\mathbb{E}\{\mathbf{x}[n]\mathbf{x}^\top[n]\}\mathbb{E}\{\mathbf{a}[n]\} \\ &= \left(\mathbf{I} - \mu\mathbf{R}\right)\mathbb{E}\{\mathbf{a}[n]\} \end{aligned} \tag{1.17}$$

where, since the principle of orthogonality, which states that the optimal estimation error and the input data samples to a Wiener filter are orthogonal (uncorrelated), has been exploited. In particular, the third step of (1.17) exploits the independence assumption, which asserts that the present observation samples $\mathbf{x}[n]$ are independent from the past observations $\mathbf{x}[n-1], \mathbf{x}[n-2]$,... This assumption allows us to write $\mathbb{E}\{\mathbf{x}[n]\mathbf{x}^\top[n]\mathbf{a}[n]\} = \mathbb{E}\{\mathbf{x}[n]\mathbf{x}^\top[n]\}\mathbb{E}\{\mathbf{a}[n]\}$. Using an eigenvalue decomposition, the

auto-correlation matrix $\mathbf{R}$ can be written as[4]

$$\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\top} \tag{1.18}$$

where $\mathbf{Q}$ is a square unitary matrix whose columns are the orthonormal eigenvectors of $\mathbf{R}$ and $\mathbf{\Lambda}$ is a diagonal matrix which contains the corresponding eigenvalues $\lambda_0, \lambda_1, ..., \lambda_{N-1}$. Now, substituting (1.18) into (1.17), leads to

$$\begin{aligned}\mathbb{E}\{\mathbf{a}[n+1]\} &= \left(\mathbf{I} - \mu\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\top}\right)\mathbb{E}\{\mathbf{a}[n]\} \\ &= \left(\mathbf{Q}\mathbf{Q}^{\top} - \mu\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\top}\right)\mathbb{E}\{\mathbf{a}[n]\} \\ &= \mathbf{Q}\left(\mathbf{I} - \mu\mathbf{\Lambda}\right)\mathbf{Q}^{\top}\mathbb{E}\{\mathbf{a}[n]\} \ . \end{aligned} \tag{1.19}$$

By left-multiplying both sides of (1.19) by $\mathbf{Q}^{\top}$, one gets

$$\mathbf{Q}^{\top}\mathbb{E}\{\mathbf{a}[n+1]\} = \underbrace{\mathbf{Q}^{\top}\mathbf{Q}}_{\mathbf{I}}\left(\mathbf{I} - \mu\mathbf{\Lambda}\right)\mathbf{Q}^{\top}\mathbb{E}\{\mathbf{a}[n]\}$$

which can be simplified as

$$\mathbf{a}'[n+1] = \left(\mathbf{I} - \mu\mathbf{\Lambda}\right)\mathbf{a}'[n] \tag{1.20}$$

having defined $\mathbf{a}'[n] \triangleq \mathbf{Q}^{\top}\mathbb{E}\{\mathbf{a}[n]\}$. This vector $\mathbf{a}'[n]$ represents the average error weight vector $\mathbf{a}[n]$, but for a rotation introduced by the unitary matrix $\mathbf{Q}^{\top}$. Since a rotation preserves the amplitude, the convergence of the average error vector in terms of $\mathbf{a}'[n]$ can be studied.

In a scalar form, (1.20) becomes

$$a'_i[n+1] = \left(1 - \mu\lambda_i\right)a'_i[n] \qquad i = 0, 1, ..., N-1 \ . \tag{1.21}$$

[4]Since $\mathbf{R}$ is a correlation matrix, it is symmetric and non negative definite. As a consequence, its eigenvalues are non negative and its eigenvectors can be selected as orthonormal. Under these conditions, the eigenvalue decomposition (1.18) is also interpretable as a special case of Singular Value Decomposition (SVD) [17].

Starting with initial values $a'_0[0], a'_1[0], ..., a'_{N-1}[0]$, (1.21) can be expressed as

$$\begin{aligned} a'_i[1] &= (1 - \mu\lambda_i)\, a'_i[0] \\ a'_i[2] &= (1 - \mu\lambda_i)\, a'_i[1] \\ &= (1 - \mu\lambda_i)\,(1 - \mu\lambda_i)\, a'_i[0] \\ &= (1 - \mu\lambda_i)^2\, a'_i[0] \end{aligned}$$

which, in general, becomes

$$a'_i[n] = (1 - \mu\lambda_i)^n\, a'_i[0], \qquad i = 0, 1, ..., N-1\ . \tag{1.22}$$

Equation (1.22) implies that $\mathbf{a}'[n]$ convergences to zero if and only if

$$|1 - \mu\lambda_i| < 1, \qquad i = 0, 1, ..., N-1$$

or, equivalently,

$$0 < \mu < \frac{2}{\lambda_i}, \qquad i = 0, 1, ..., N-1\ .$$

This leads to a constraint for the step-size parameter μ to guarantee LMS convergence as

$$0 < \mu < \frac{2}{\lambda_{\max}}.$$

Normalized LMS

Besides having a convergence that depends on the step-size parameter, the standard LMS algorithm is sensitive to a scaling of its input. To cope with this issue, a variant referred to as Normalized LMS algorithm, can be derived. Let us consider the following tap weight update recursion

$$\mathbf{w}[n+1] = \mathbf{w}[n] + \mu[n]e[n]\mathbf{x}[n] \tag{1.23}$$

which differs from (1.15) only in the fact that the step-size parameter $\mu[n]$ is now time-varying [19]. In order to derive a time variant step-size, the key idea is to consider the one-step forward error defined with respect to the tap weights at the next time step

$n+1$ as:

$$e^+[n] = d[n] - \mathbf{w}^\top[n+1]\mathbf{x}[n] \ . \tag{1.24}$$

By substituting (1.23) in (1.24), one has

$$\begin{aligned} e^+[n] &= d[n] - \left(\mathbf{w}[n] + \mu[n]e[n]\mathbf{x}[n]\right)^\top \mathbf{x}[n] \\ &= d[n] - \mathbf{w}^\top[n]\mathbf{x}[n] - \mu[n]e[n]\mathbf{x}^\top[n]\mathbf{x}[n] \\ &= \ e[n] - \mu[n]e[n]\mathbf{x}^\top[n]\mathbf{x}[n] \\ &= \left(1 - \mu[n]\mathbf{x}^\top[n]\mathbf{x}[n]\right) e[n] \ . \end{aligned}$$

Minimizing $\left(e^+[n]\right)^2$ with respect to $\mu[n]$ yields

$$\begin{aligned} \frac{\partial \left(e^+[n]\right)^2}{\partial \mu[n]} &= \frac{\partial}{\partial \mu[n]} \left(1 - \mu[n]\mathbf{x}^\top[n]\mathbf{x}[n]\right)^2 \left(e[n]\right)^2 \\ &= -2\left(1 - \mu[n]\mathbf{x}^\top[n]\mathbf{x}[n]\right)\left(\mathbf{x}^\top[n]\mathbf{x}[n]\right)\left(e[n]\right)^2 \ . \end{aligned}$$

Equating to zero and solving with respect to $\mu[n]$, one obtain

$$\mu[n] = \frac{1}{\mathbf{x}^\top[n]\mathbf{x}[n]}$$

which gives us the step-size value which minimizes the mean square value of the one-step forward error $e^+[n]$ at time step n. Equation (1.23) can be, therefore, rewritten as

$$\mathbf{w}[n+1] = \mathbf{w}[n] + \frac{1}{\mathbf{x}^\top[n]\mathbf{x}[n]} e[n]\mathbf{x}[n]$$

in which the step-size parameter is selected to be proportional to the inverse of the instantaneous signal energy $\mathbf{x}^\top[n]\mathbf{x}[n]$. However note that, if the instantaneous signal energy becomes very small, the fraction diverges. In order to cope with this issue, the following modified recursion is usually adopted

$$\boxed{\mathbf{w}[n+1] = \mathbf{w}[n] + \tilde{\mu}\frac{\mathbf{x}[n]}{\alpha + \mathbf{x}^\top[n]\mathbf{x}[n]} e[n] \ .} \tag{1.25}$$

being $\tilde{\mu}$ the new step-size parameter which controls the convergence and α a positive constant, small enough when compared with $\mathbf{x}^\top[n]\mathbf{x}[n]$.

Leaky Normalized LMS

The last LMS-based method here described is called *Leaky Normalized* LMS algorithm. This method is the one used in this project. As previously mentioned, since the cost function is quadratic, the LMS algorithm solves a convex problem. From a geometrical point of view, the LMS solution is a recursive approach to find the vertex position of a paraboloid, by moving step by step in the space of the tap weights, starting from the past position.

Under certain conditions, the updated filter coefficients $\mathbf{w}[n+1]$ should exhibit less correlation with the past ones. In this case, a parameter that weights the past filter coefficient values, named *leakage factor*, is introduced in the LMS strategy. In this sense, the leakage factor allows to weight the past differently, e.g., for $\lambda < 1$ the tap weights at time $n+1$ have lower dependence on the past step n. The leaky normalized LMS update recursion is written as

$$\boxed{\mathbf{w}[n+1] = \lambda \mathbf{w}[n] + \tilde{\mu} \frac{\mathbf{x}[n]}{\alpha + \mathbf{x}^{\top}[n]\mathbf{x}[n]} e[n] \ .} \tag{1.26}$$

where $\lambda \in [0, 1]$ is the leakage factor. Note that for $\lambda = 1$, (1.26) reduces to (1.25). For $\lambda < 1$, the weights are updated accounting for reduced memory on their past values.

Chapter 2

Reference Scenario

In this chapter, the reference scenario and the considered setup for the simulation-based analysis are reported. The rest of this chapter is organized as follows. In Section 2.1, the simulation setup and the used measurements obtained from a car interior are presented. System performance indicators are presented in Section 2.2. Finally, in Section 2.3 the best grouping of the reference signals is shown.

2.1 Simulation Setup

A general block diagram of single reference FF SISO ANC system with FxLMS algorithm is shown in Fig. 2.1. The filter $P(z)$, the so-called *primary path*, represents the physical path between the vibration (captured by the accelerometer) and the perceived audio (captured by the error microphone). Its output $d[n]$ is usually referred to as *desired output*, since it represents the signal to be detected. Similarly, at the output of the adaptive filter $W(z)$, the so-called *anti-noise* signal $y[n]$ is given. The purpose of audio signal $y[n]$, emitted by the loudspeaker, is to cope with the interior noise of the cabin perceived by the error microphone. However, we have to take into account the physical path between the loudspeaker and the error microphone. This is referred to as *secondary path* and is denoted by the block $S(z)$. Hence, at the error microphone, the acoustic sum of signals $d[n]$ and $y'[n]$ is obtained.

A detailed mathematical derivation of a multiple reference FF MIMO ANC system with FxLMS algorithm is presented in Section 3.1.

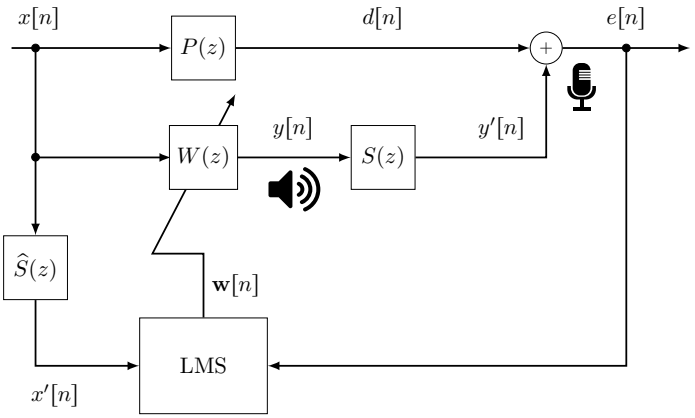


Figure 2.1: Block diagram of single reference adaptive FF SISO ANC system.

The considered signals come from an experimental measurement campaign performed by ASK on a realistic car interior, i.e., that of a sedan (class D). In the following, J, K and M denote the number of reference signals, loudspeakers and microphones respectively. All signals are acquired with a sampling frequency of 48 kHz and down-sampled to 3 kHz for a time duration of about 82.22 seconds, i.e., 246679 samples. Six reference signals ($J = 6$), i.e. $\mathbf{x}_j[n] = [x_j[n], x_j[n-1], ..., x_j[n-N+1]]^\top$ for $j = 1, 2, ..., J$, were recorded by accelerometers fixed at crucial positions of the car structure, obtaining thus a multi-reference scenario. In particular, the considered crucial positions are the front and rear car body and front suspensions. Since the identification of the desired output improves if more than one reference is available, it is expected that the greater the number of reference signals, the larger the noise cancellation effect. The number of possible combinations of references is

$$\sum_{k=1}^{6} \binom{6}{k} = 6 + 15 + 20 + 15 + 6 + 1 = 63$$

since only 6 combinations are possible to arrange the six reference signals one by one, 15 are possible to arrange two references at a time (1-2, 1-3, 1-4,...,5-6), and so on. Two ($M = 2$) error microphones are considered. They are binaural microphones placed by the driver's left and right ears to capture the perceived audio. Two loudspeakers ($K = 2$) are taken into account; this yields four (KM) possible combinations of secondary paths when a full MIMO ANC system is considered. Accordingly, a secondary path

matrix can be defined as

$$\mathbf{S}(z) = \begin{bmatrix} S_{11}(z) & S_{12}(z) \\ S_{21}(z) & S_{22}(z) \end{bmatrix} \begin{matrix} \} \ \mathbf{S}_1(z) \\ \} \ \mathbf{S}_2(z) \end{matrix}$$

where the first subscript denotes the error microphone and the second one identifies the loudspeaker, so that the row vectors $\mathbf{S}_1(z)$ and $\mathbf{S}_2(z)$ identify the secondary paths related to the first and the second error microphone, respectively.

In order to analyze the behavior of the considered secondary paths, the magnitude frequency responses, the impulse responses, and the group delays against frequency of these four filters are shown in Fig. 2.2 for $S_{11}(z)$ and $S_{12}(z)$ and in Fig. 2.3 for $S_{21}(z)$ and $S_{22}(z)$. In particular, the identification of the secondary paths is performed by using an adaptive approach, i.e., by means the LMS algorithm, with white noise as input signal. More details on the adopted estimation method can be found in Section 6.4 (see Fig. 6.13). From these figures, we can observe that, in the frequency-domain, secondary paths are low-pass filters well approximated with about 240 taps and a gain in the frequency range going from 20 Hz to the cut-off frequency about 520 Hz. The band-pass is therefore about 500 Hz and in this bandwidth there may be a strong distortion. By observing the impulse response of $S_{11}(z)$ we can note that the filter has an accumulate delay of about 6–7 ms, since the first peak occurs after 18–22 samples. In this context, the group delay together with the impulse response give us information on the time of flight of the audio signal. In these figures, the insets show the group delay in the band-pass [0.2 0.52] kHz. For example, in Fig. 2.2, first column relative to microphone 1 and loudspeaker 1, the group delay is about 25 samples, in agreement with the impulse response. For the sake of clarity, Tabs. 2.1 and 2.2 summarize the adopted nomenclature and the values of the LMS parameters used in our simulations, respectively. Simulations are performed by means of properly designed Simulink-based models and scripts.

2.2 Performance Indicators

The obtained numerical results are compared in terms of the following performance measures:

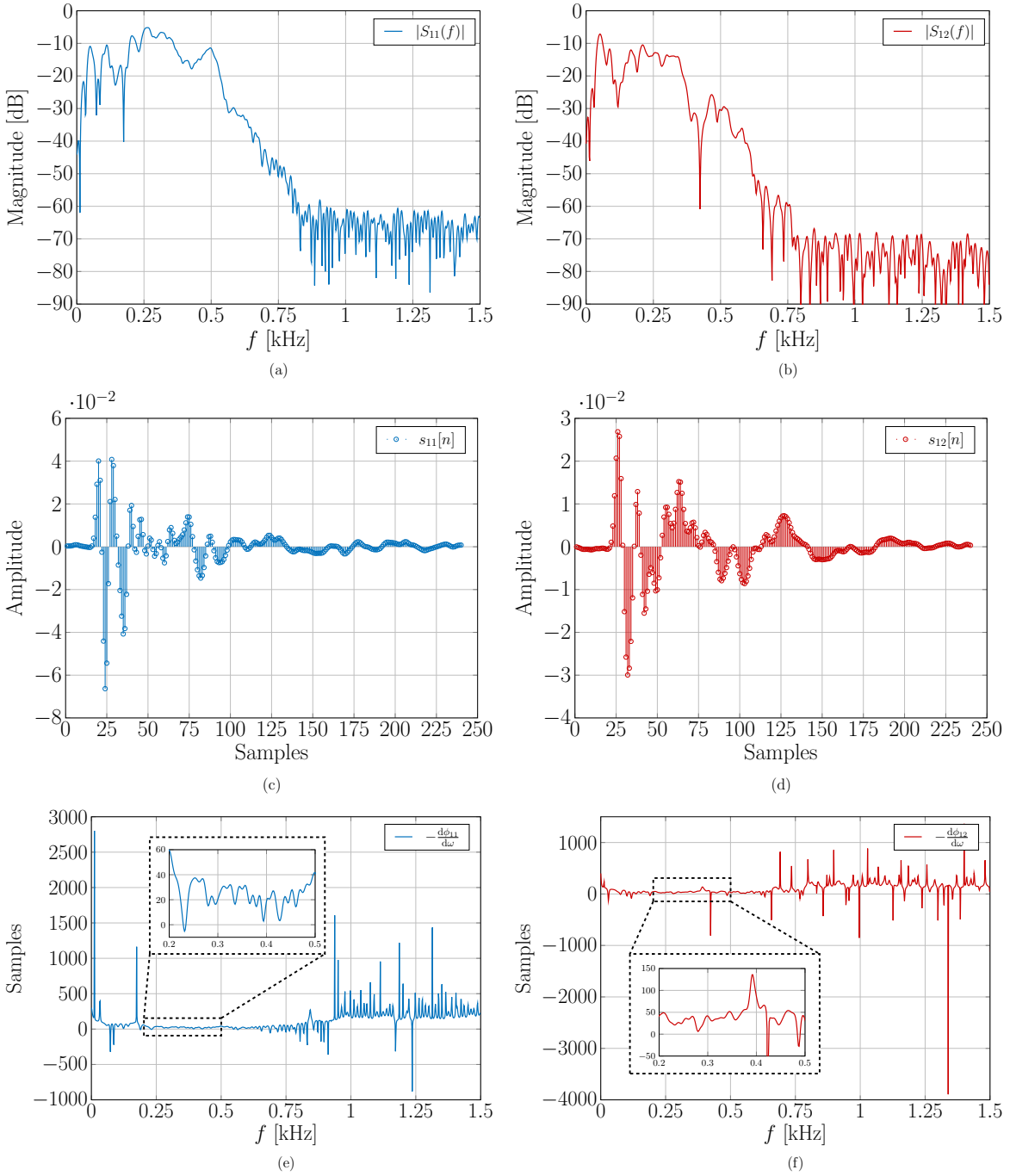


Figure 2.2: Left side: magnitude frequency response (a), impulse response (c) and group delay against frequency (e) of the secondary path between speaker 1 and left error microphone. Right side: magnitude frequency response (b), impulse response (d) and group delay against frequency (f) of the secondary path between speaker 2 and left error microphone.

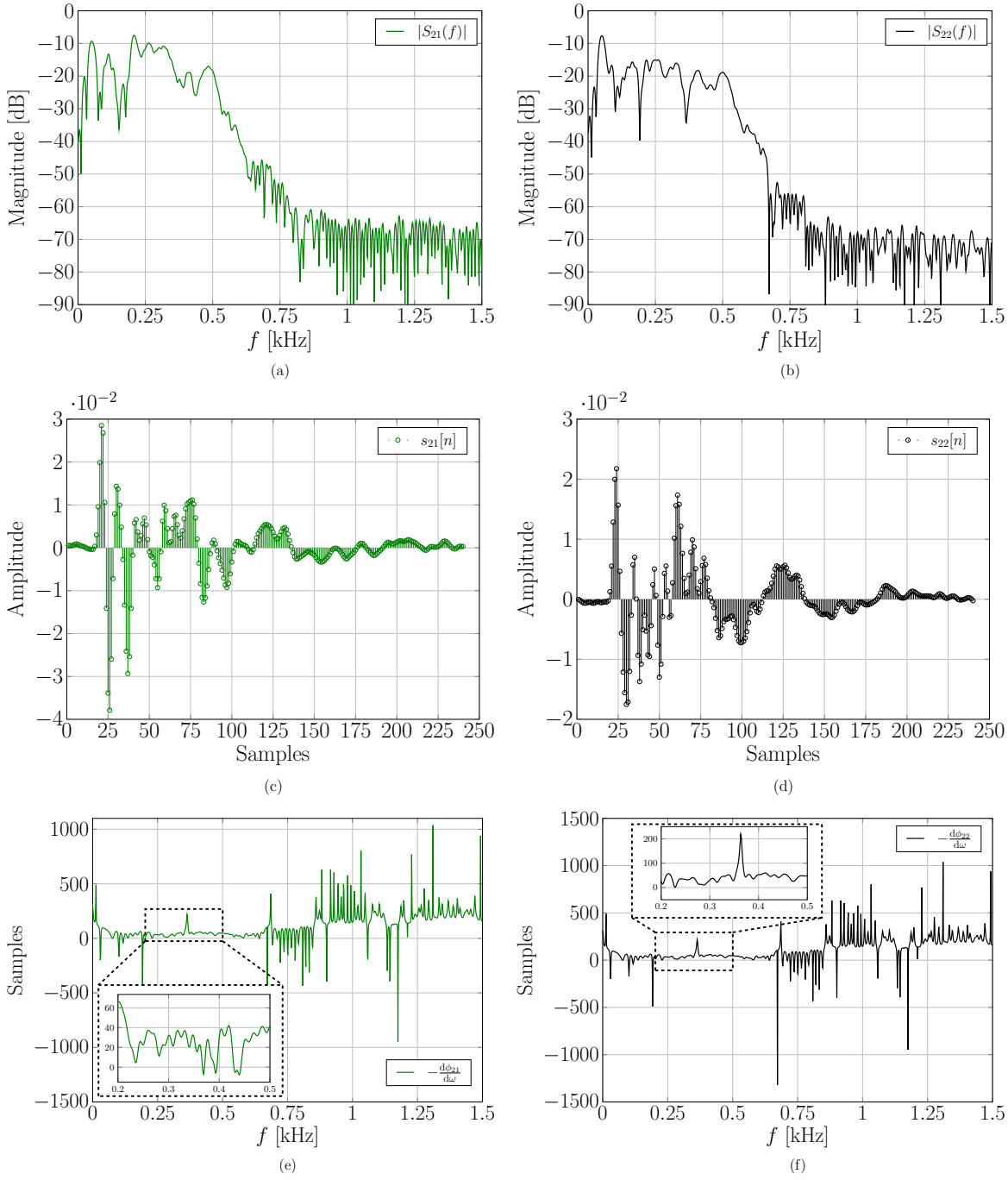


Figure 2.3: Left side: magnitude frequency response (a), impulse response (c) and group delay against frequency (e) of the secondary path between speaker 1 and right error microphone. Right side: magnitude frequency response (b), impulse response (d) and group delay against frequency (f) of the secondary path between speaker 2 and right error microphone.

Table 2.1: Adopted nomenclature.

Nomenclature	Description
J	Number of reference signals
K	Number of loudspeakers
M	Number of error microphones
$\mathbf{x}_j[n]$	Column vector of j-th reference signal
$d_m[n]$	m-th desired output
$e_m[n]$	m-th error microphone signal
$y_k[n]$	k-th anti-noise signal
$y'_m[n]$	m-th output signal

Table 2.2: LMS parameters.

Parameter	Value
Algorithm type	Leaky Normalized LMS
Number of taps	100
Step-size parameter	$\tilde{\mu} = 10^{-4}$
Leakage factor	$\lambda = 1$

- Average SPL and spectrogram of disturbance signal $d[n]$ and error $e[n]$.
- Spectral coherence $C_{\mathbf{x},d}$ between reference signals $\{x_i[n]\}$ and microphone one.
- Effective cancellation $\Gamma(f)$.

These parameters are detailed in the following subsections.

Sound Pressure Level

To understand the acoustic impact of the ANC system, it is of interest to determine the SPL of the acoustic signal at the error microphone (acquired in Volts). The SPL or acoustic pressure level is a measurement of the effective sound pressure of a mechanical wave with respect to a reference sound source [20]. A logarithmic measure of the SPL

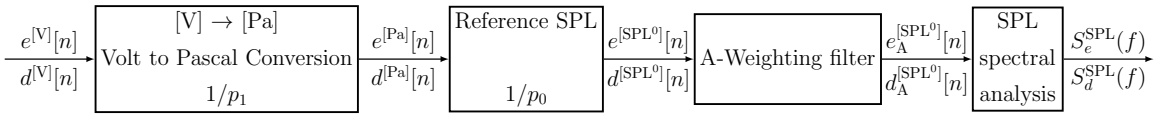


Figure 2.4: Block diagram for SPL spectral analysis.

spectrum is given by

$$S_\star^{\mathrm{SPL}}(f) = 20\log_{10}\left(\frac{p(f)}{p_0}\right) \quad [\mathrm{dB}] \tag{2.1}$$

where $\star$ can be either $d[n]$ or $e[n]$, $p(f)$ is the sound pressure in a specified Resolution BandWidth (RBW) centered at frequency f, measured in μPa, and p_0 is the reference sound pressure [20]. The RBW used in this analysis is 6 Hz. The most common reference sound pressure in air is $p_0 = 20$ μPa which is considered as the threshold of human hearing [21]. Figure 2.4 depicts a block diagram representing the conversion of both error signal and desired output. Signals in Volts, namely $e^{[\mathrm{V}]}[n]$ and $d^{[\mathrm{V}]}[n]$, are firstly scaled by a conversion factor p_1 that takes into account to the sensitivity of the microphone and then by p_0. At the output of the "Reference SPL" block, signals $e^{[\mathrm{SPL}^0]}[n]$ and $d^{[\mathrm{SPL}^0]}[n]$ are passed through the A-weighting filter. The A-weighting filter accounts for the perceived sensitivity of the human ear [22]. The sound pressure level weighted by the A-weighting filter is measured in dBA/Hz. Finally the block "SPL spectral analysis" estimates the relevant SPL spectrum values according to (2.1). The time-averaged SPL spectrum can be calculated by the pspectrum MATLAB function which evaluates the power spectrum $S_\star^{\mathrm{SPL}}(f)$.

Furthermore, pspectrum allows to evaluate the spectrogram of the audio signal. The spectrogram is a visualization of the intensity of a sound as a function of frequency and time. It is usually depicted as a heat map [23] by varying the color according to the increase/decrease of the power spectrum.

Spectral Coherence

The spectral coherence between a set of reference signals $\mathbf{X}[n] = [\mathbf{x}_1[n], \mathbf{x}_2[n], ..., \mathbf{x}_j[n]]^\top$ and desired output $d[n]$ gives us information on how much the signal $d[n]$ is correlated with the reference signals. As a consequence, it measures the amount of noise that can be canceled and, more important, in which frequency range the cancellation is

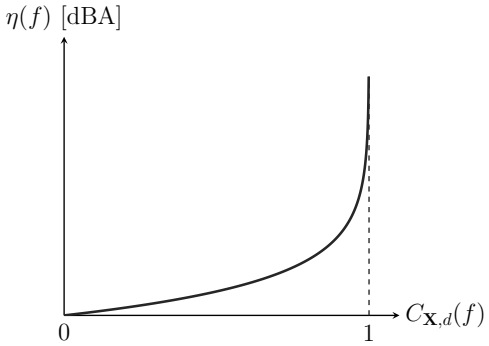


Figure 2.5: Theoretical cancellation $\eta(f)$ as a function of spectral coherence $C_{\mathbf{X},d}(f)$.

expected to be effective. The spectral coherence is defined as [24]

$$C_{\mathbf{X},d}(f) = \frac{P^{\dagger}_{\mathbf{X}d}(f) P^{-1}_{\mathbf{XX}}(f) P_{\mathbf{X}d}(f)}{P_{dd}(f)} \tag{2.2}$$

where, $P_{\mathbf{XX}}(f)$ and $P_{dd}(f)$ are the power spectral densities of the matrix input signals $\mathbf{X}[n]$ and $d[n]$, respectively, $P_{\mathbf{X}d}(f)$ is their cross power spectral density, and † denotes the hermitian operator. Note that $C_{\mathbf{X},d}(f) \in [0, 1]$, i.e., if $C_{\mathbf{X},d}(f) = 0$ the signals are uncorrelated, whereas if $C_{\mathbf{X},d}(f) = 1$ they are highly correlated. These operations can be performed by the MATLAB function mscohere.

This measure can be used to evaluate the theoretical cancellation $\eta(f)$ that gives us an upper bound on the noise cancellation. The theoretical cancellation is equal to [25]

$$\eta(f) \triangleq -10 \log_{10} \left[1 - C_{\mathbf{X},d}(f)\right] \quad \text{[dBA]} \ . \tag{2.3}$$

The idea behind (2.3) is that if $C_{\mathbf{X},d}(f) = 0$, $\eta(f)$ reduces to zero, and there is no cancellation. Conversely, if $C_{\mathbf{X},d}(f) = 1$ (the signals are perfectly correlated), $\eta(f) \to +\infty$, obtaining a theoretically infinite cancellation. Fig. 2.5 shows the behavior of the theoretical cancellation $\eta(f)$ versus the spectral coherence $C_{\mathbf{X},d}(f)$ according to (2.3).

Effective Cancellation

Since the auditory perception is extremely subjective, a main task of an ANC system is to find a fair performance indicator. To this end, we define the effective cancellation $\Gamma(f)$ as the difference between the average SPL spectrum of the desired signal $S_d^{\mathrm{SPL}}(f)$

and that of the error signal $S_e^{\mathrm{SPL}}(f)$, i.e.,

$$\Gamma(f) \triangleq S_d^{\mathrm{SPL}}(f) - S_e^{\mathrm{SPL}}(f) \qquad [\mathrm{dBA}] \ . \tag{2.4}$$

If $\Gamma(f) < 0$, the ANC system is worsening the performance since the anti-noise power is higher than the power of the desired output. Conversely, if $\Gamma(f) > 0$ the ANC system is working properly, since it is canceling the desired output.

Alternatively, in certain circumstances, as we shall see later, it may be useful to analyze this parameter in the linear domain. In this case, the parameter can be evaluated as

$$\Gamma_{\mathrm{lin}}(f) \triangleq \frac{10^{\frac{S_d^{\mathrm{SPL}}(f)}{20}}}{10^{\frac{S_e^{\mathrm{SPL}}(f)}{20}}} \ .$$

The effective cancellation $\Gamma(f)$ in (2.4) can be compared with the theoretical cancellation expressed $\eta(f)$ in (2.3) [25].

2.3 Best Reference Signals Grouping

As previously stated, we have 63 possible combinations for arranging the six reference signals in order to get good performance. In our analysis, the best grouping for a given value of J, i.e., $J = 1, 2, 3, ..., 6$, is obtained for the maximum average cancellation. The average cancellation is defined as

$$\overline{\Gamma} \triangleq 20 \log_{10} \left(\frac{1}{\gamma_f} \int_{f_{\min}}^{f_{\max}} \Gamma_{\mathrm{lin}}(f) \, df \right) \tag{2.5}$$

where γ_f is the measure of the frequency interval of our interest defined as $\gamma_f = f_{\max} - f_{\min}$, being, in our case, $f_{\min} = 20$ Hz and $f_{\max} = 520$ Hz. Suppose that for $J = 2$ the reference signals which show the best average of $\Gamma(f)$ were references x_1, x_3; then we write the best grouping as $G_2 : [x_1, x_3]$. For the same G_2, we also evaluate the absolute maximum value of $\Gamma(f)$ and the corresponding frequency as an additional performance measure.

A different approach was adopted for the evaluation of the minimum. In fact, for each grouping, we firstly pick up the contiguous frequencies in which $\Gamma(f)$ experiences the largest number of local minima and then we average such local minimum values.

Let us suppose that for $J = 2$ the reference signals which show the largest number of contiguous minima of $\Gamma(f)$ in the frequency range $\gamma_{f,\min}$ were references x_2, x_6; then we write the worst grouping as $W_2 : [x_2, x_6]$.

For W_2 we evaluate a "coherent" average of the local minima in the corresponding $\gamma_{f,\min}$ and consider only that grouping and value as the minimum. This approach allows us to discard negative spikes in a specific frequency range and consider a larger frequency range in which then ANC system is worsening the performance.

Figure 2.6 shows an example which summarizes the evaluation of average, absolute maximum and minimum of $\Gamma(f)$. In particular, the green curve depicts $\Gamma(f)$ for the best grouping G_2, and the dashed blue line depicts its average value $\overline{\Gamma}$. The absolute maximum $\Gamma_{\max} = 6.82$ dBA is obtained at 208.5 Hz and the absolute minimum $\Gamma_{\min} = -0.42$ dBA at 137.75 Hz. The orange curve shows $\Gamma(f)$ in case of the "worst" grouping W_2. The negative lobes ($\Gamma(f) < 0$) for both curves are highlighted in gray.

Only the widest negative lobe relative to W_2 is considered for the evaluation of the minimum. Within $\gamma_{f,\min} = [432, 520]$ Hz, the average of SPL is evaluated as follows

$$\overline{\Gamma}_{\min} = 20 \log_{10} \left(\frac{1}{\gamma_{f,\min}} \int_{f \in \gamma_{f,\min}} \Gamma_{\text{lin}}(f)\, df \right) .$$

Note that our analysis is based on absolute frequencies. Since they depend on the considered frequency resolution, this analysis may be extended to sub-band of frequencies, as proposed in [26].

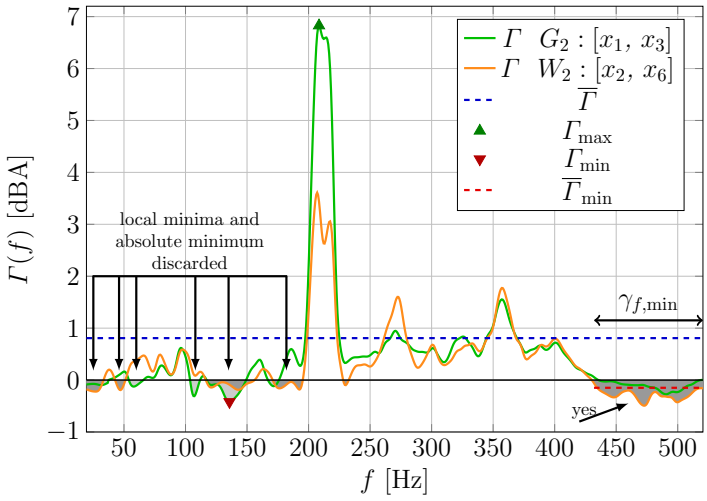


Figure 2.6: Example of average, absolute maximum and minimum evaluation for finding the best grouping of reference signals.

Chapter 3

FeedForward Schemes

In this chapter, the main results on multiple reference FF ANC schemes are presented. The chapter is organized as follows. For the sake of compactness, the mathematical background is given only for the multiple reference MIMO FF ANC system in Section 3.1. Hence, the SISO and MISO schemes can be considered as special cases of the MIMO one, having $K = M = 1$ and $K \neq 1, M = 1$, respectively. Numerical results are presented in Sections 3.2.1, 3.2.2 and 3.2.3 for the SISO, MISO and MIMO ANC system, respectively

3.1 Multiple Reference MIMO System

In this section, the adaptive FF MIMO ANC scheme description is given. This system, depicted in Fig. 3.1, is referred to as FxLMS for the reasons explained in the following discussion.

The physical channels between the vibrations captured by the J accelerometers and the audio waves perceived by the error microphones, can be represented by FIR filters with transfer function $\mathbf{P}(z)$. At the m-th error microphone, the signal $d_m[n]$[5] can be described as a sum of discrete convolutions between the J reference signals $\{x_j[n]\}$ and

[5]The disturbing signal is usually referred to as *desired output*, since it represents the signal to be detected.

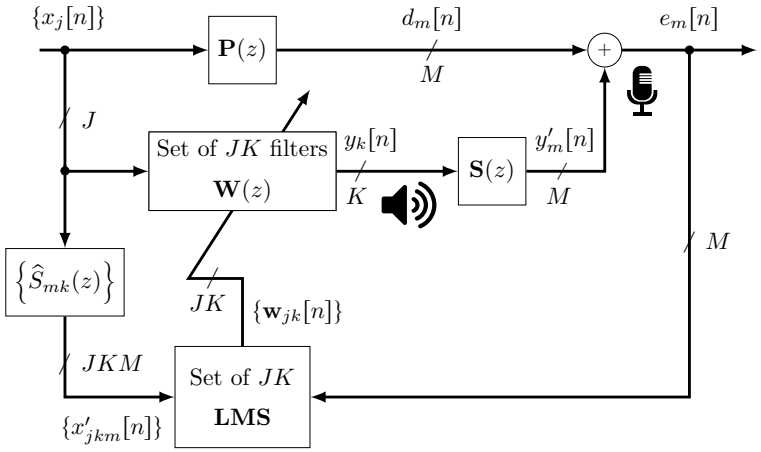


Figure 3.1: Discrete-time block diagram of multiple reference FF MIMO ANC system with FxLMS algorithm.

the Q-length filter impulse responses, as

$$d_m[n] = \sum_{j=1}^{J} \sum_{q=0}^{Q-1} x_j[n-q] p_{jm}[q] = \sum_{j=1}^{J} \mathbf{x}_j^\top[n] \mathbf{p}_{jm}$$

where $\mathbf{p}_{jm} = [p_{jm}[0], p_{jm}[1], ..., p_{jm}[Q-1]]^\top$ represents the impulse response vector from the j-th reference signal to the m-th disturbing signal, with $\mathcal{Z}$-transform specified by the jm-term of the matrix response filter $\mathbf{P}(z)$ of size $J \times M$ and $\top$ denotes the transpose operator. Similarly, the output of the N-length adaptive filter $\mathbf{W}(z)$,[6] the *anti-noise* signal emitted by the k-th loudspeaker $y_k[n]$ is obtained as

$$y_k[n] = \sum_{j=1}^{J} \sum_{\ell=0}^{N-1} x_j[n-\ell] w_{jk\ell}[n] = \sum_{j=1}^{J} \mathbf{w}_{jk}^\top[n] \mathbf{x}_j[n] \tag{3.1}$$

where

$$\mathbf{x}_j[n] = [x_j[n], x_j[n-1], ..., x_j[n-N+1]]^\top \tag{3.2}$$

and the control filter coefficient vector is defined as

$$\mathbf{w}_{jk}[n] = [w_{jk0}[n], w_{jk1}[n], w_{jk0}[n] ..., w_{jk,N-1}[n]]^\top \tag{3.3}$$

[6]Since this adaptive filter is time varying, the notation $\mathbf{W}(z)$ is misleading. We retain this notation under the assumption of slow time variation.

which denotes the impulse response vector at the n-th time-epoch of the adaptive filter from the j-th reference signal to the k-th loudspeaker signal. The aim of the anti-noise signal $y_k[n]$, emitted by the k-th loudspeaker, is to cope with the interior noise of the cabin perceived by the error microphones. The physical acoustic channels between loudspeakers and error microphones are referred to as *secondary paths* and can be described by set of low-pass FIR filters $\mathbf{S}(z)$. Hence, the signal $y'_m[n]$ is obtained by the sum of all the anti-noise signals emitted by the K loudspeakers and filtered by the corresponding secondary paths. Mathematically, one gets

$$\begin{aligned} y'_m[n] =& \sum_{k=1}^{K} y_k[n] \otimes s_{mk}[n] \\ =& \sum_{k=1}^{K} \sum_{h=0}^{H-1} s_{mk}[h] y_k[n-h] \\ =& \sum_{k=1}^{K} \sum_{h=0}^{H-1} s_{mk}[h] \sum_{j=1}^{J} \sum_{\ell=0}^{N-1} x_j[n-\ell-h] w_{jk\ell}[n-h] \end{aligned} \tag{3.4}$$

where $s_{mk}[n]$ identifies the impulse response of the secondary path $S_{mk}(z)$ from the k-th loudspeaker to the m-th error microphone of length H and $\otimes$ denotes the convolution operator. Finally, at the m-th error microphone, the acoustic sum of the disturbing audio wave $d_m[n]$ and its noise-cancellation signal $y'_m[n]$ is obtained, as

$$e_m[n] = d_m[n] + y'_m[n] \ . \tag{3.5}$$

The aim of the LMS algorithm is to estimate the tap-weights of the control filter in order to minimize a suitable cost function in the mean square sense. Since the LMS input is a filtered version of the reference signals, this algorithm is known in the literature as Filtered-X-LMS. For a MIMO system, the relevant cost function to be minimized is defined as the sum of the squared error signals coming from the M error microphones. Mathematically, one obtains

$$\xi[n] = \sum_{m=1}^{M} e_m^2[n] \ . \tag{3.6}$$

The gradient of (3.6) with respect to the proper tap-weights is calculated as

$$\nabla\xi[n] = \nabla \sum_{m=1}^{M} e_m^2[n] = 2 \sum_{m=1}^{M} e_m[n] \nabla e_m[n] \ . \tag{3.7}$$

The elements of the gradient of the m-th error microphone signals are

$$\begin{aligned}
\frac{\partial e_m[n]}{\partial w_{jki}} &= \frac{\partial}{\partial w_{jki}} \left(d_m[n] + y'_m[n] \right) \\
&= \frac{\partial}{\partial w_{jki}} \left(\sum_{k'=1}^{K} \sum_{h=0}^{H-1} s_{mk'}[h] y_{k'}[n-h] \right) \\
&= \frac{\partial}{\partial w_{jki}} \left[\sum_{k'=1}^{K} \sum_{h=0}^{H-1} s_{mk'}[h] \left(\sum_{j'=1}^{J} \sum_{\ell=0}^{N-1} x_{j'}[n-\ell-h] w_{j'k'\ell} \right) \right] \\
&= s_{mk}[n] \otimes x_j[n-i] \\
&= x'_{jkm}[n]
\end{aligned} \tag{3.8}$$

where under the assumption of perfect secondary path estimation, i.e., $\widehat{S}_{mk}(z) = S_{mk}(z)$, $x'_{jkm}[n-i] = x_j[n-i] \otimes \hat{s}_{mk}[n] = x_j[n-i] \otimes s_{mk}[n]$ for $m = 1, 2, ..., M$ and $i = 0, 1, ..., N-1$. By defining

$$\begin{aligned}
\mathbf{x}'_{jkm}[n] &= \left[x'_{jkm}[n], x'_{jkm}[n-1], ..., x'_{jkm}[n-N+1] \right]^\top \\
&= \left[\hat{s}_{mk}[n] \otimes x_j[n], \hat{s}_{mk}[n] \otimes x_j[n-1], ..., \hat{s}_{mk}[n] \otimes x_j[n-N+1] \right]^\top
\end{aligned}$$

and substituting (3.8) into (3.7) it is possible to compactly write

$$\nabla\xi[n] = 2 \sum_{m=1}^{M} e_m[n] \mathbf{x}'_{jkm}[n]$$

for $j = 1, 2, ..., J$ and $k = 1, 2, ..., K$. Now, remembering that the steepest descent algorithm is defined as (1.10), the tap-weight recursion of the leaky normalized LMS algorithm for a MIMO system is

$$\boxed{\mathbf{w}_{jk}[n+1] = \lambda \mathbf{w}_{jk}[n] - \tilde{\mu} \sum_{m=1}^{M} e_m[n] \frac{\mathbf{x}'_{jkm}[n]}{\alpha + \mathbf{x}'^{\top}_{jkm}[n] \mathbf{x}'_{jkm}[n]} \ .} \tag{3.9}$$

Unlike the weight-update equations in Chapter 1, the update term in (3.9), has a minus sign because the error (3.5) is now defined with a plus sign.

Further mathematical background on FF ANC system can be found in Appendix A.

3.2 Numerical Results

In this section, numerical results on the FF ANC schemes are presented. For the sake of compactness, only the left error microphone is here considered. In particular, for all schemes, only the best grouping with one, three, five and six references, i.e., $J = 1, 3, 5, 6$, are considered according to the procedure described in Section 2.3. Further results can be found in Appendix B.

3.2.1 SISO

The SISO ANC system employs one error microphone and one loudspeaker for the anti-noise emission. Hence, only the secondary path S_{11} described in Section 2.1 is employed in these simulations.

Fig. 3.2 shows the SPL spectra as a function of frequency for the best grouping with $J = 1$ (a), $J = 3$ (b), $J = 5$ (c) and $J = 6$ (d). It is possible to observe that the SPL values of the green curve, which denotes the sum of the interior car noise plus the anti-noise, is almost always below the audio signal level of the interior car noise shown by the red curve, regardless the number of reference signals. This shows that noise cancellation is effective. Moreover, the higher the number of active reference signals, the better the performance. In fact, the best performance occurs for $J = 6$. As expected, the noise cancellation is effective within the range 200–450 Hz.

Figure 3.3 shows the average $\overline{\Gamma}$ (2.5), the maximum $\Gamma_{\max}$ and minimum $\Gamma_{\min}$ of effective cancellation $\Gamma(f)$ (2.4) in dBA for the best grouping. The average and maximum of $\Gamma(f)$ increase according to J, whereas the minimum of $\Gamma(f)$ decreases with J. In particular, the FF ANC system experiences a peak cancellation $\Gamma_{\max}$ around 208–210 Hz. This frequency range can be associated with the noise caused by the cavity modes of the tire rolling on the asphalt [5, 25]. The frequency range in which we have the smallest cancellation, i.e., the smallest $\Gamma_{\min}$ goes from 435 to 513 Hz. In particular, with more than 3 reference signals, the SISO system approximately yields

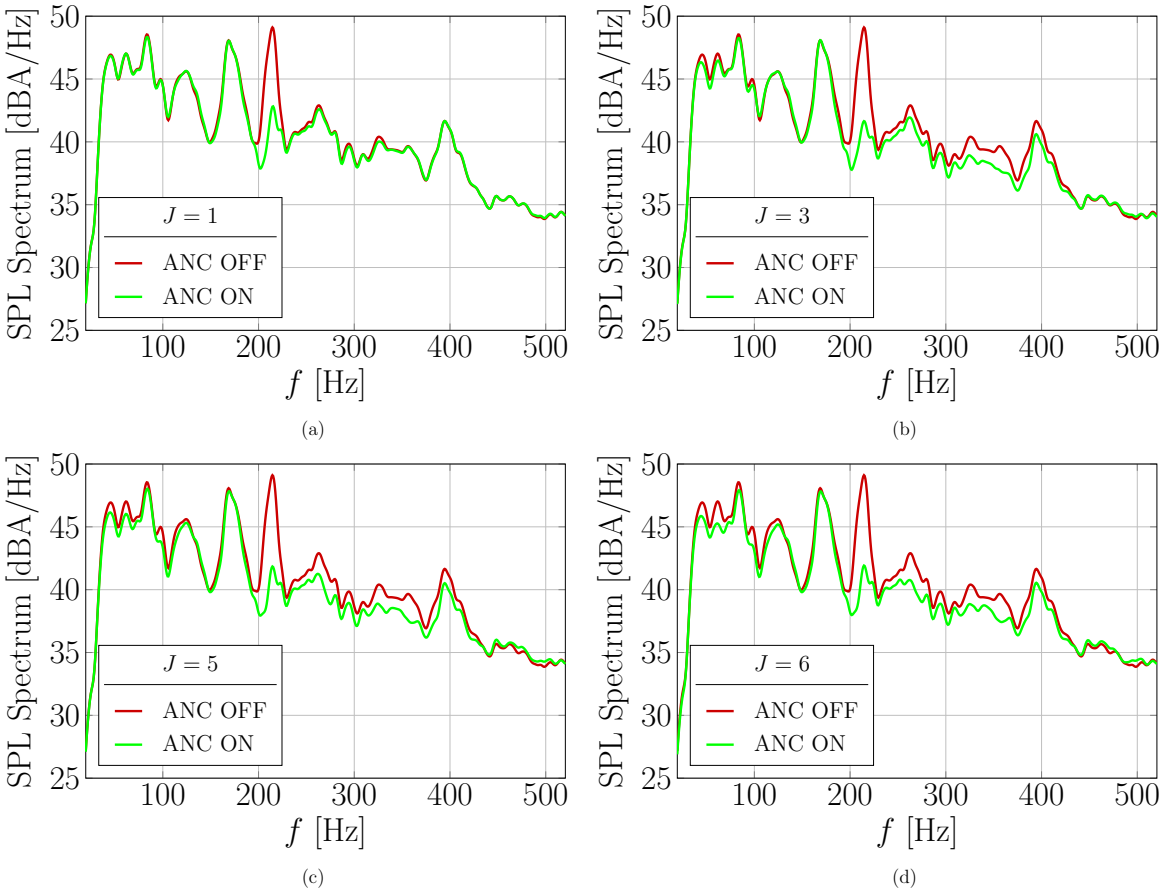


Figure 3.2: SPL spectra of $d[n]$ and $e[n]$ in case of SISO ANC system for the best grouping of reference signals with $J = 1$ (a), $J = 3$ (b), $J = 5$ (c) and $J = 6$ (d).

the same minimum in this frequency range.

In Fig. 3.4, the effective cancellation $\Gamma(f)$ (2.4) and theoretical cancellation $\eta(f)$ are shown, respectively, as functions of frequency, for the best groupings with $J = 1$ (a), $J = 3$ (b), $J = 5$ (c) and $J = 6$ (d). The larger the number of available reference signals, the greater the amount of noisy spectrum that can be canceled. This is noticeable by observing how the theoretical cancellation improves. The gap between theoretical and effective cancellation is due to the presence of the secondary path. Although this gap in some frequency range is remarkable, it is important to note how at about 210 Hz the two curves are quite similar to each other. This means that, even in the presence of the secondary path, we are able to cancel the noise in this frequency range almost as much as in theory. Finally, in Fig. 3.5 the spectrogram is shown both in the case of ANC

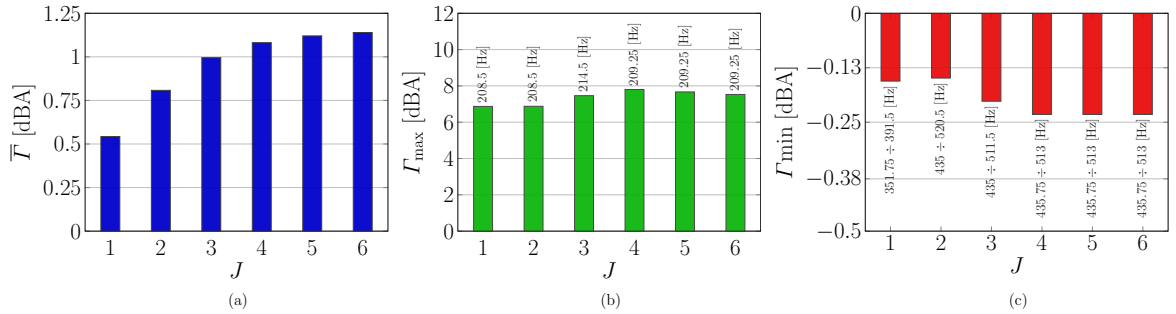


Figure 3.3: Average (a), maximum (b) and minimum (c) cancellation $\Gamma(f)$ as functions of J in case of SISO ANC system.

system ON for $J = 6$ and OFF. It is easy to observe how the noise components around 200–250 Hz are perfectly canceled during the whole simulation time. Furthermore, it is appreciable how the noise is somewhat attenuated also in other frequency ranges, e.g., 50–150 Hz and 400–410 Hz, after the first 35 seconds of simulation.

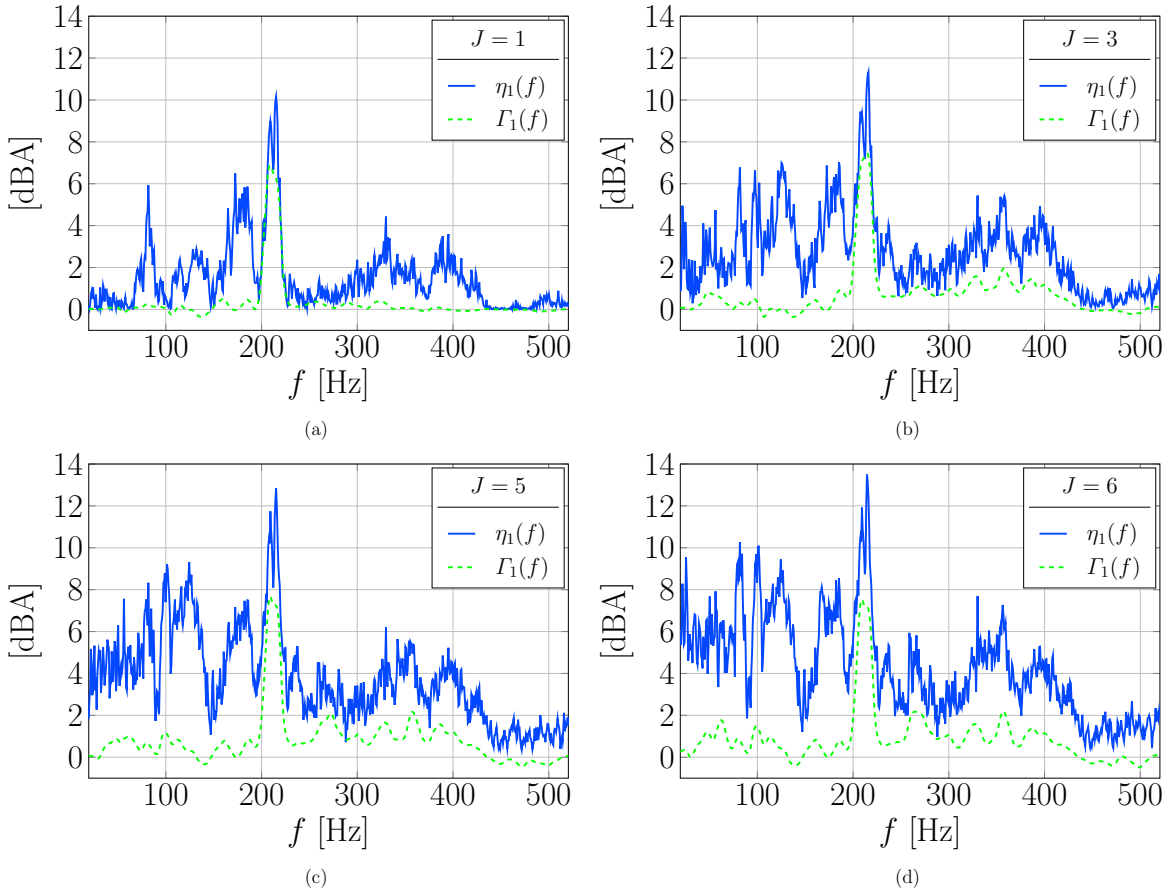


Figure 3.4: Comparison between theoretical cancellation $\eta(f)$ and effective cancellation $\Gamma(f)$, as a function of frequency, in case of SISO ANC system for the best grouping of reference signals with $J = 1$ (a), $J = 3$ (b), $J = 5$ (c) and $J = 6$ (d).

3.2.2 MISO

The spectrum of the desired output (in red) and error signal (in green) are shown in Fig. 3.6 for $J = 1$ (a), $J = 3$ (b), $J = 5$ (c) and $J = 6$ (d) for the case of MISO FF ANC system with 2 loudspeakers and 1 microphone, i.e., $K = 2$ and $M = 1$. The SPL spectrum values of the error signal decrease for increasing number of reference signals, showing a good noise cancellation. Note that, in this case, just three reference signals are sufficient to achieve the best cancellation. Figure 3.7 shows the main measures of $\Gamma(f)$, namely the average $\overline{\Gamma}$, absolute maximum $\Gamma_{\max}$ and minimum $\Gamma_{\min}$ and corresponding frequencies. It can be observed that, even if the absolute maximum is achieved employing only two reference signals, the average performance

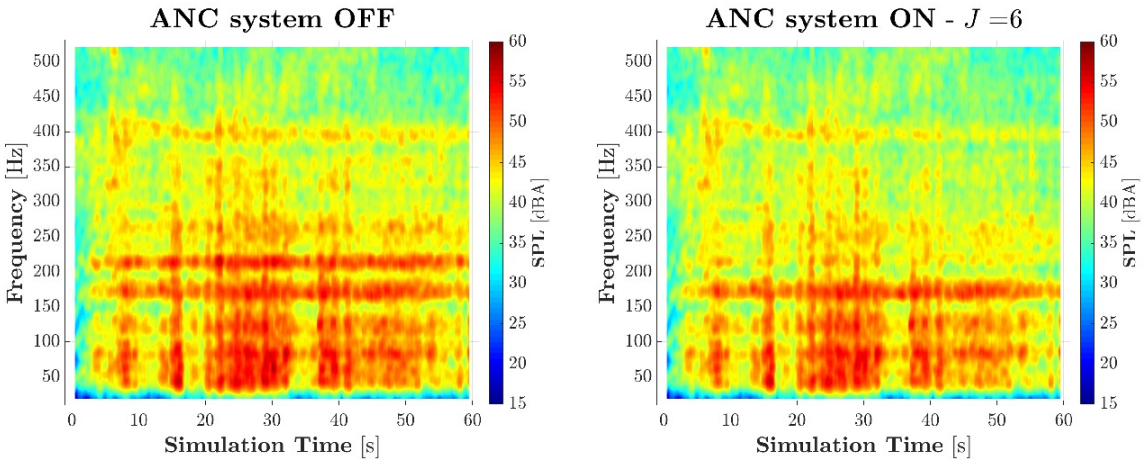


Figure 3.5: Spectrogram of the error microphone signal when the SISO FF ANC system is switched OFF (left) and ON (right) for $J = 6$.

increases for increasing number of the references, as expected. Furthermore note that by using few signal references, in this case, the minimum is concentrated around 300–350 Hz, whereas it occurs in the range 430–520 Hz for larger number of reference signals.

The comparison between theoretical and effective cancellation $\eta(f)$ and $\Gamma(f)$ as a function of frequency is shown in Fig. 3.8 for $J = 1$ (a), $J = 3$ (b), $J = 5$ (c) and $J = 6$ (d). In agreement with the above discussion, it is possible to observe that, for $J = 3$, $\Gamma(f)$ almost fits $\eta(f)$ in the frequency. The spectrogram plots are depicted in Fig. 3.9. It is possible to observe that, similarly to the SISO case, the main canceled noise contributions are around the tire cavity mode frequency range [1,5]. However, by employing the MISO system, we are able to cancel a wider noise spectrum compared to the SISO case.

The corresponding spectrogram analysis of the left error microphone signal when the MISO FF ANC system is switched OFF (left) and ON (right) with $J = 6$ reference signals is shown in Fig. 3.9.

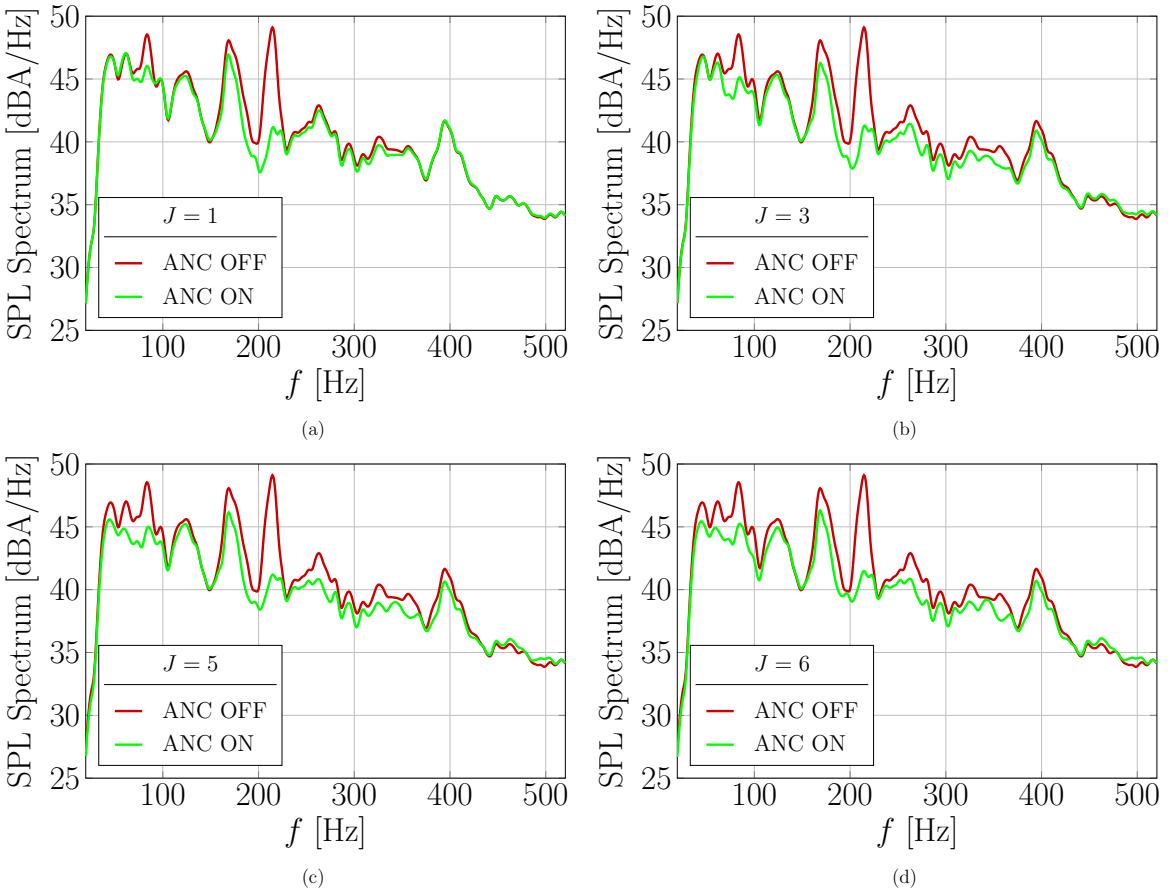


Figure 3.6: SPL spectra of $d[n]$ and $e[n]$ in case of MISO ANC system for the best grouping of reference signals with $J = 1$ (a), $J = 3$ (b), $J = 5$ (c) and $J = 6$ (d).

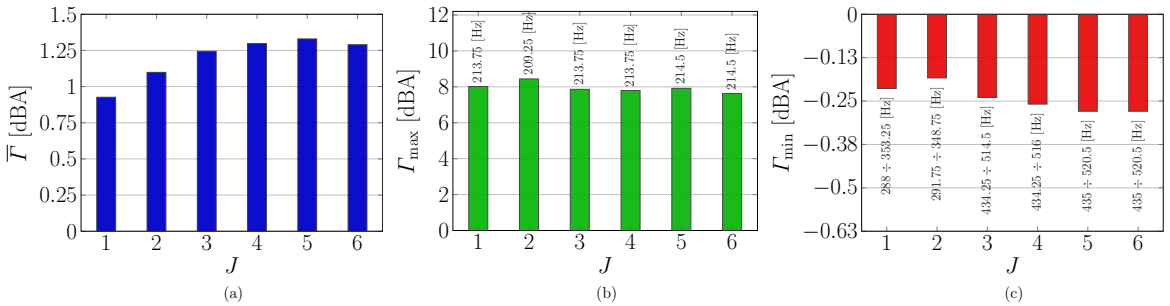


Figure 3.7: Average (a), maximum (b) and minimum (c) cancellation $\Gamma(f)$ as functions of J in case of MISO ANC system.

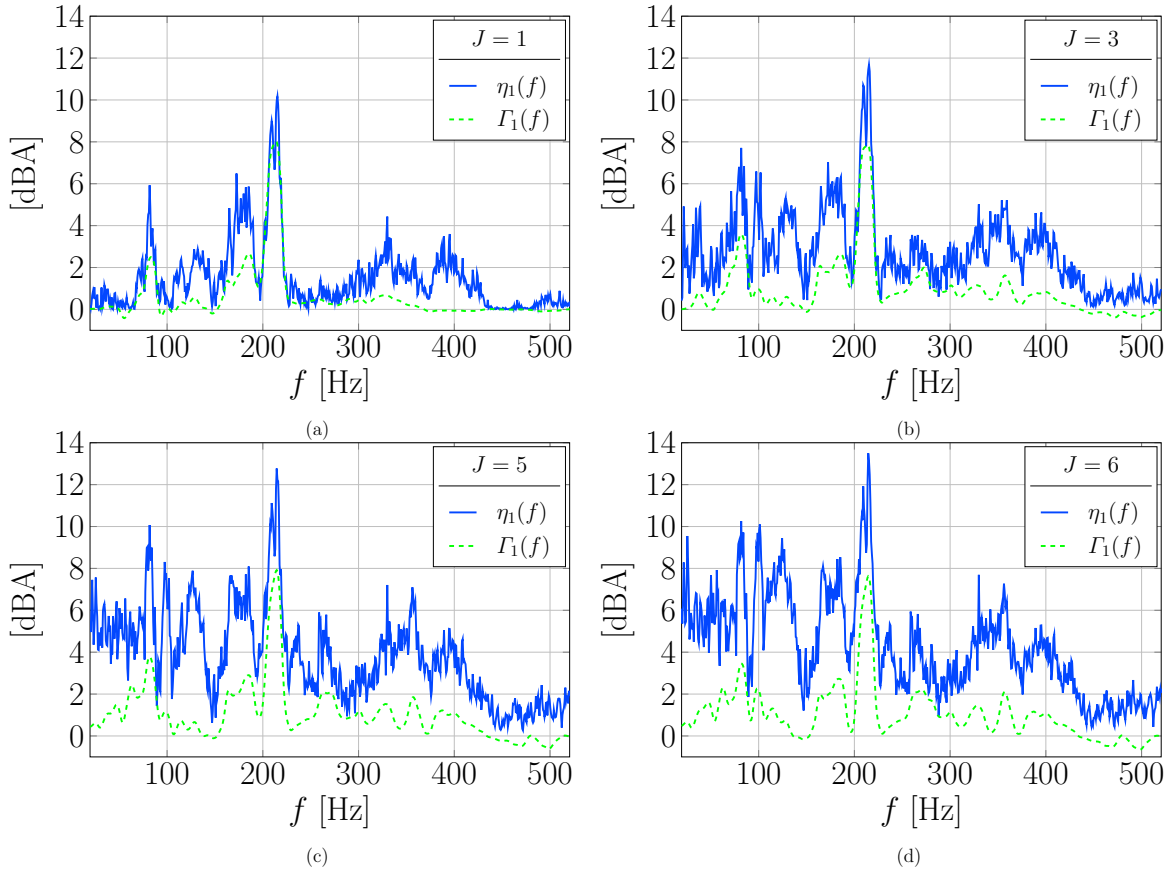


Figure 3.8: Comparison between theoretical cancellation $\eta(f)$ and effective cancellation $\Gamma(f)$, as a function of frequency, in case of MISO ANC system for the best grouping of reference signals with $J = 1$ (a), $J = 3$ (b), $J = 5$ (c) and $J = 6$ (d).

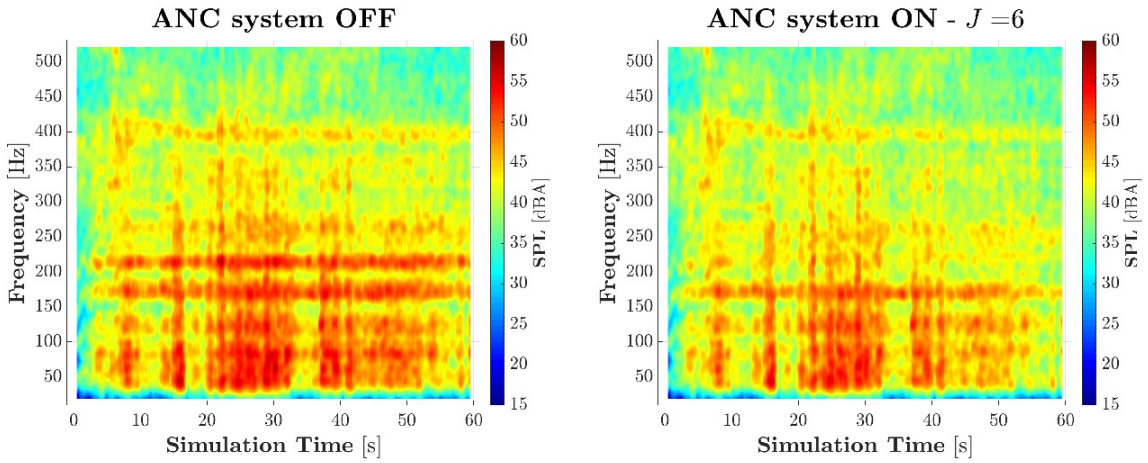


Figure 3.9: Spectrogram of the error microphone signal when the MISO FF ANC system is switched OFF (left) and ON (right) for $J = 6$.

3.2.3 MIMO

In Figs. 3.10 and 3.11 the SPL spectra and the theoretical/effective cancellation comparison for a MIMO ANC system, i.e., by employing $K = M = 2$, for the best grouping of reference signals with $J = 1$ (a), $J = 3$ (b), $J = 5$ (c) and $J = 6$ (d) is shown, respectively. One can observe that the low frequencies, i.e., 20–120 Hz are well mitigated. A maximum of cancellation, in the range 200–220 Hz is exhibited.

For both error microphones, in Fig. 3.12 the main measures, i.e., average $\overline{\Gamma}$ (a), maximum $\Gamma_{\max}$ (b) and minimum $\Gamma_{\min}$ (c), of the effective cancellation $\Gamma(f)$ for the best grouping and the corresponding frequencies are shown. Note that the left error microphone has, on the average, better performance than the right one. Contrary to the previous cases, by increasing the number of signal references no improvements are exhibited. For both microphones, the absolute maximum points are set around 200–220 Hz. The minima are concentrated within 300–400 Hz regardless of the considered error microphone.

For the sake of completeness, the spectrogram plots for left error microphone when the MIMO ANC system is OFF and ON by using $J = 6$ reference signals, are shown in Fig. 3.13. Note that the cancellation peak is always present during the whole simulation time.

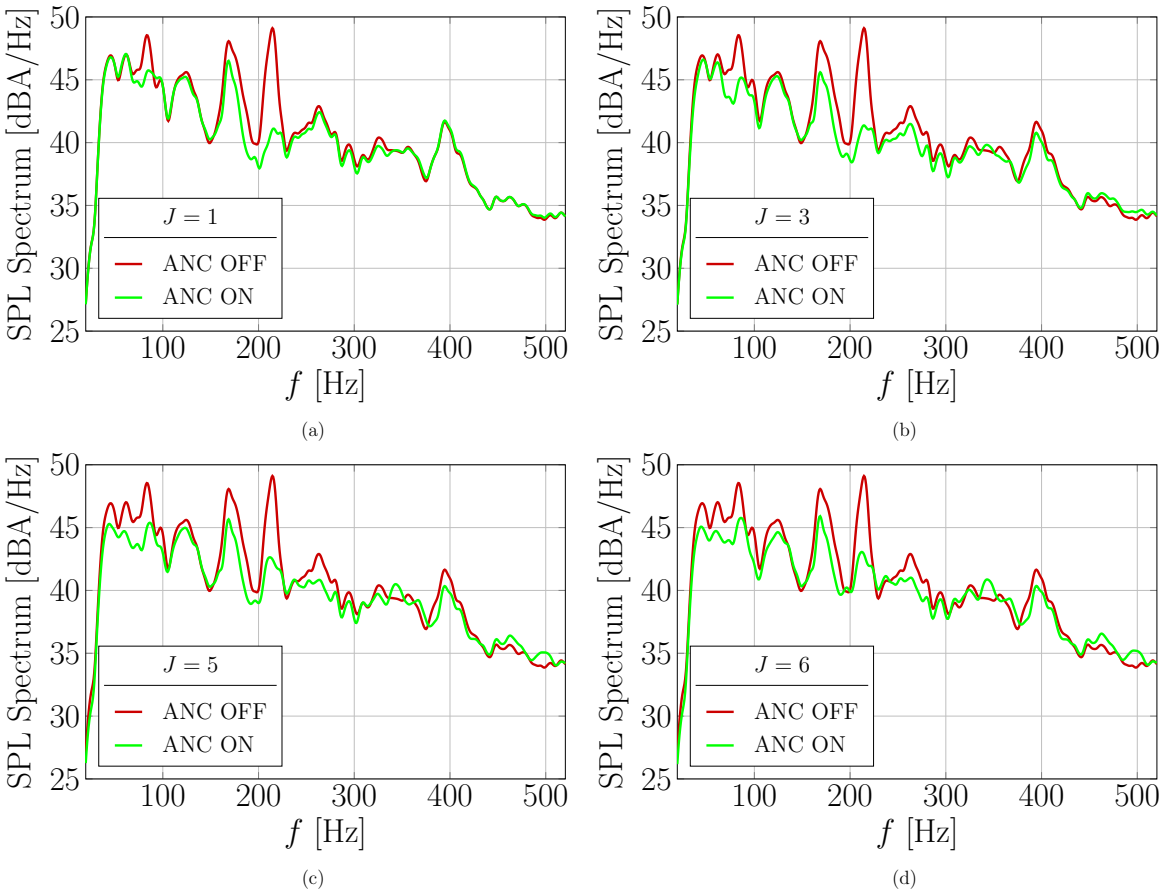


Figure 3.10: SPL spectra of $d[n]$ and $e[n]$ in case of MIMO ANC system for the best grouping of reference signals with $J = 1$ (a), $J = 3$ (b), $J = 5$ (c) and $J = 6$ (d).

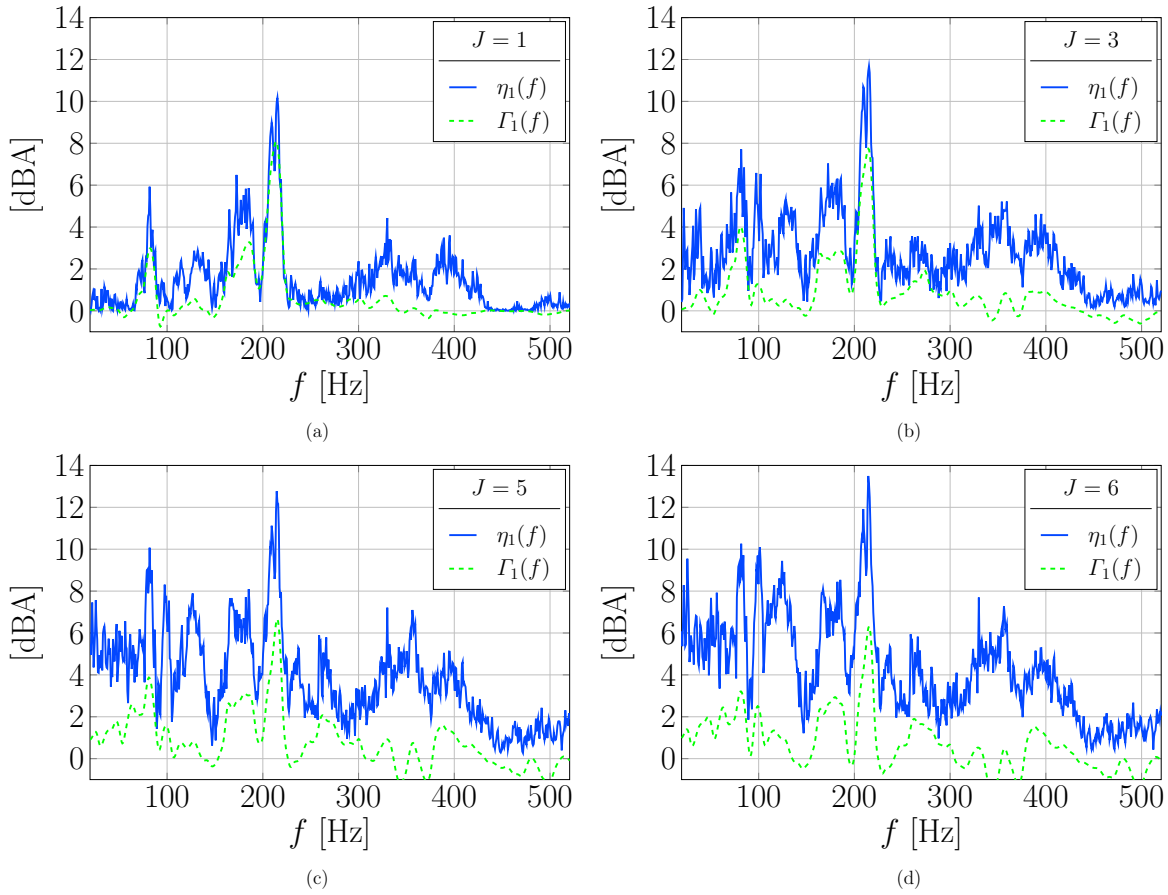


Figure 3.11: Comparison between theoretical cancellation $\eta(f)$ and effective cancellation $\Gamma(f)$, as a function of frequency, in case of MIMO ANC system for the best grouping of reference signals with $J = 1$ (a), $J = 3$ (b), $J = 5$ (c) and $J = 6$ (d).

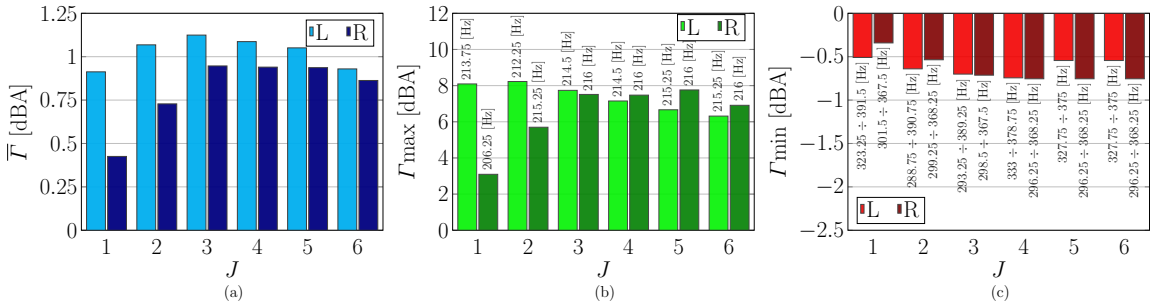


Figure 3.12: Average (a), maximum (b) and minimum (c) cancellation $\Gamma(f)$ as functions of J for both left (L) and right (R) error microphones in case of MIMO ANC system.

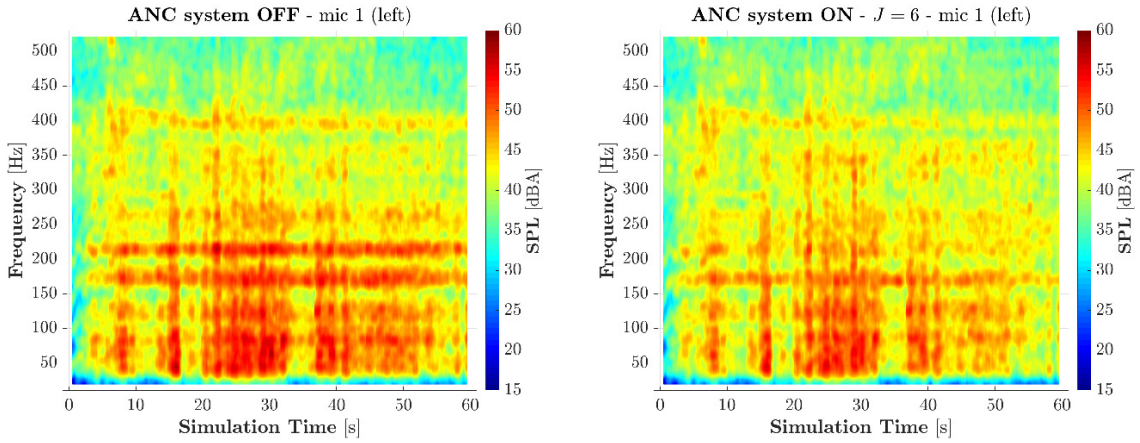


Figure 3.13: Spectrogram of the left error microphone signal when the MIMO FF ANC system is switched OFF (left) and ON (right) for $J = 6$.

Chapter 4

FeedBack Schemes

The first part of this report is dedicated to the implementation and analysis of FF ANC systems. It is shown that the prior knowledge of the reference signals, acquired by accelerometers positioned in the car structure, allows significant benefits in terms of noise cancellation. However, it is known that when reference signals are unavailable or they are incoherent with respect to the error microphone signals, the FF ANC system is ineffective [27,28]. For this reason, it may be preferable to employ a FB ANC system. In fact, due to the lack of reference signals, the FB ANC system tries to cancel the disturbing audio noise in the overall frequency spectrum.

It is possible to distinguish FB ANC systems based on the algorithm implemented in the feedback loop; in particular, the adaptive approaches such as FxLMS or Modified FxLMS or fixed ones based on control theory. The general FB MIMO ANC system is shown in Fig. 4.1. For adaptive approaches, the idea is to synthesize reference signal(s) by using the error microphone signal(s). In this sense, the FB ANC schemes differ with respect to the FF ones only in the absence of primary path(s) and reference signals. In particular, $d_m[n]$ refers to the m-th desired output, $e_m[n]$ to the m-th error microphone signal, $S_{mk}(z)$ to the secondary path from the k-th loudspeaker to the m-th error microphone and filter $\mathcal{K}_{mk}(z)$ indicates a generic feedback controller which operates under a certain algorithm. In this chapter, several FB schemes implementing different filters $\mathcal{K}_{mk}(z)$ are discussed.

The rest of this chapter is organized as follows. The mathematical derivation of adaptive FB ANC systems is presented in Section 4.1. In Section 4.2, a fixed controller

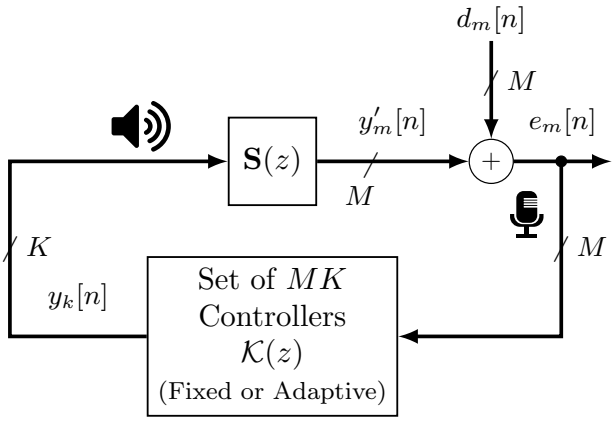


Figure 4.1: General block diagram of the FB MIMO ANC system.

for the SISO FB ANC system is proposed.

4.1 Adaptive FB Systems

In this section, a description of the considered adaptive FB ANC systems is given. Similarly to the FF ANC system, in Section 4.1.1, the "standard" FxLMS approach is derived for a FB scheme, whereas, in Section 4.1.2, a different version of this algorithm, known as Modified-FxLMS or Delayless-FxLMS, is presented. Note that, for the sake of generality, in this section only MIMO systems, both for standard and modified LMS, are considered.

In Appendix E further numerical results on adaptive FB ANC systems can be found. Moreover, in Section E.3, a special case of MIMO FB system, the so-called FB Combined Input - Combined Output (CICO) system is analyzed.

4.1.1 Standard FxLMS

The block diagram of the standard FxLMS scheme [3] for an adaptive FB MIMO ANC system is depicted in Fig. 4.2, in which the controller $\boldsymbol{\mathcal{K}}(z)$ has been highlighted. Following the literature [3], the disturbance signals $d_m[n]$ are referred to as *desired outputs* since they represent the signals to be detected. The error microphone signal $e_m[n]$ is expressed as in (3.5). Similarly, at the m-th error microphone, signal $y'_m[n]$ is

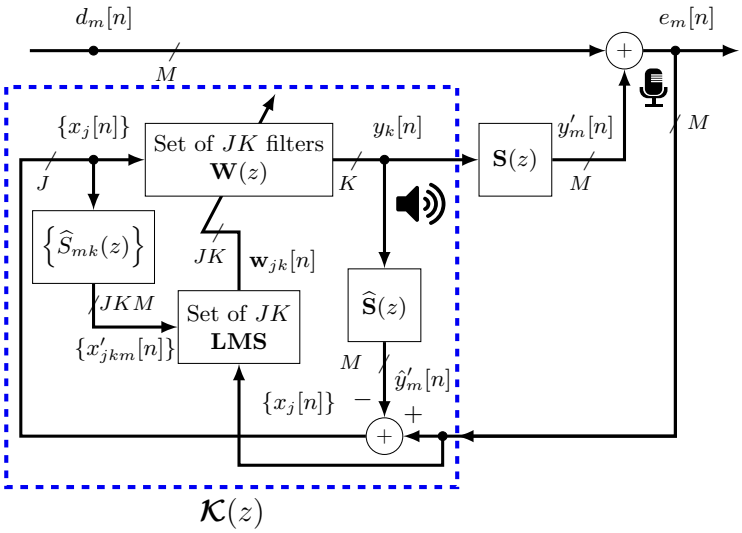


Figure 4.2: Block diagram of the standard FxLMS for an adaptive FB MIMO ANC system.

expressed in the definition of (3.4), i.e.,

$$y'_m[n] = \sum_{k=1}^{K} y_k[n] \otimes s_{mk}[n] \quad m = 1, 2, ..., M \ . \tag{4.1}$$

In Fig. 4.2, the block $\mathbf{S}(z)$ performs the operation (4.1) with elements $S_{mk}(z)$. Unlike the FF system, where reference signals are available, here they have to be estimated. By inverting (3.5), a set of reference signals $\{x_j[n]\}$, as close as possible to the noise sources, can be synthesized. The block $\widehat{\mathbf{S}}(z)$ performs an identical operation with elements $\widehat{S}_{mk}(z)$, which denote estimates of $S_{mk}(z)$, whereas the block $\{\widehat{S}_{mk}(z)\}$ performs an element-wise filtering operation in which each input signal $x_j[n]$ is filtered by $\widehat{S}_{mk}(z)$ yielding the output $x'_{jkm}[n]$, as defined in the following derivation.

For reference signals derivation starting from the error microphone ones, let us invert (3.5), such that the desired output is expressed as

$$d_m[n] = e_m[n] - y'_m[n] \ .$$

The goal is to synthesize a reference signal $x_m[n]$ as close as possible to $d_m[n]$. This

can be done by estimating the desired output as

$$x_m[n] = e_m[n] - \hat{y}'_m[n] \quad m = 1, 2, ..., M \tag{4.2}$$

where at the m-th error microphone the estimated anti-noise signal at the output of the estimated secondary path $\hat{y}'_m[n]$ is defined as

$$\hat{y}'_m[n] = \sum_{k=1}^{K} y_k[n] \otimes \hat{s}_{mk}[n]$$

where $\hat{s}_{mk}[n]$ is the impulse response of the estimated filter $\widehat{S}_{mk}(z)$. One can note that the main difference between the FB MIMO scheme of Fig. 4.2 and the FF MIMO scheme depicted in Fig. 3.1 is the presence of the blocks $\left\{\widehat{S}_{mk}(z)\right\}$ and $\widehat{\mathbf{S}}(z)$, based on the same elements, one used for FxLMS algorithm (as in FF system) and one for reference signal reconstruction purposes.

The input signal of the $k-$th loudspeaker $y_k[n]$ can be written as

$$y_k[n] = \sum_{j=1}^{M} \sum_{\ell=0}^{N-1} x_j[n-\ell] w_{jk\ell}[n] = \sum_{j=1}^{M} \mathbf{x}_j[n] \mathbf{w}_{jk}^{\top}[n] \tag{4.3}$$

having defined the reference signal vector $\mathbf{x}_j[n]$ and the impulse response at the $n-$th time epoch of the adaptive filter from the $j-$th reference signal to the $k-$th loudspeaker $\mathbf{w}_{jk}[n]$ as in (3.2) and (3.3), respectively. Note that, unlike the FF ANC system, where the number of available reference signals was J, here it is constrained by the number of error microphones, i.e., $J = M$.

A limit of this system is given by the delay introduced by the secondary paths. In fact, if the secondary paths introduce too large delays in the system, the reconstruction of the reference signals $x_j[n]$ exhibits a lag with respect to the current $d_m[n]$ which may compromise the noise cancellation considerably. For this reason, it is preferable to employ a FB ANC system in case of tonal inputs, e.g., the noise produced by the car engine, where a lag of a multiple of the period is not effective.

Once the reference signals $x_m[n]$ are given, the update weight equation of the LMS algorithm follows the same mathematical steps already given in Section 3.1 for the case of a FF system. Hence, the update equation for a leaky normalized LMS FB

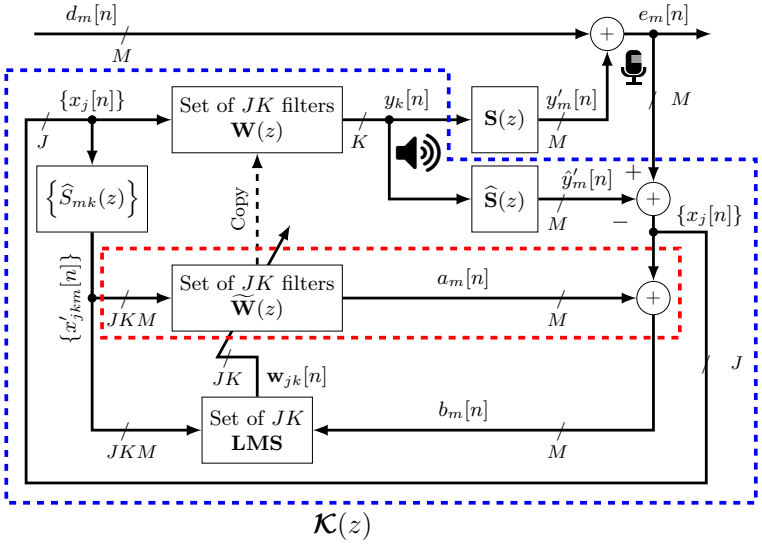


Figure 4.3: Block diagram of the modified FxLMS for adaptive FB MIMO ANC system.

MIMO ANC system, under the assumption of perfect secondary path estimation, i.e., $\hat{s}_{mk}[n] = s_{mk}[n]$, is equal to the expression given in (3.9).

4.1.2 Modified FxLMS

In this section, the mathematical derivation of the modified FxLMS scheme is reported. This scheme, proposed for the first time in [29], has been widely employed for ANC purposes because of its benefits in terms of convergence speed [5].

In Section 1.3, the LMS algorithm and its convergence is discussed, highlighting how the step-size parameter $\tilde{\mu}$ plays an important role for the convergence of the algorithm. Therefore, in order to avoid problems related to the stability of the adaptation algorithm, $\tilde{\mu}$ has to be set properly. In particular, due to the length of the system (in terms of tap weights) and the delay introduced by the secondary path, the step-size parameter is usually set to a very small number, e.g., between 10^{-4} and 10^{-5}. This leads to a slow convergence of the standard FxLMS algorithms. The aim of the Modified FxLMS is to improve the convergence speed thanks to a more accurate estimation of the error signal.

The block diagram of the modified FxLMS for a MIMO system is depicted in Fig. 4.3. By comparing the modified FxLMS scheme in Fig. 4.3 with the standard one

in Fig. 4.2, one can notice that the two schemes are essentially the same except for the presence of a second filter $\widetilde{\mathbf{W}}(z)$, which is a proper rearrangement of the top filter $\mathbf{W}(z)$, as detailed in the following. This part is highlighted in the dashed red area.

The error signal at the m-th microphone is

$$\begin{aligned} e_m[n] &= d_m[n] + y'_m[n] \\ &= d_m[n] + \sum_{k=1}^{K} y_k[n] \otimes s_{mk}[n] \\ &= d_m[n] + \sum_{k=1}^{K} s_{mk}[n] \otimes \sum_{j=1}^{M}\sum_{\ell=0}^{M} x_j[n-\ell] w_{jk\ell}[n] \quad m = 1, 2, ..., M \end{aligned}$$

where the synthesized reference signal $x_j[n]$, $j = 1, 2, ..., M$ is, expressed as in (4.2). The input of the bottom filter $\widetilde{\mathbf{W}}(z)$ is the filtered version of the reference signal $x_j[n]$ according to the theory of the FxLMS algorithm.

The output $a_m[n]$ of the bottom adaptive filter $\widetilde{\mathbf{W}}(z)$ can be expressed as

$$\begin{aligned} a_m[n] &= \sum_{k=1}^{K}\sum_{j=1}^{M}\sum_{\ell=0}^{N-1} x'_{jkm}[n-\ell] w_{jk\ell}[n] \\ &= \sum_{k=1}^{K}\sum_{h=0}^{H-1} \hat{s}_{mk}[h] \sum_{j=1}^{M}\sum_{\ell=0}^{N-1} x_j[n-\ell-h] w_{jk\ell}[n] \ . \end{aligned} \tag{4.4}$$

It is worth noting that the signal $a_m[n]$ is basically an estimate of the signal $y'_m[n]$. In fact, under the assumption of perfect secondary path estimate, i.e., $\hat{s}_{mk}[n] = s_{mk}[n]$, one obtains

$$\begin{aligned} y'_m[n] &= \sum_{k=1}^{K} s_{mk}[n] \otimes \sum_{j=1}^{M}\sum_{\ell=0}^{N-1} x_j[n-\ell] w_{jk\ell}[n] \\ &= \sum_{k=1}^{K} \hat{s}_{mk}[n] \otimes \sum_{j=1}^{M}\sum_{\ell=0}^{N-1} x_j[n-\ell] w_{jk\ell}[n] \\ &= \hat{y}'_m[n] \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{K} \sum_{j=1}^{M} \sum_{\ell=0}^{N-1} x_j[n-\ell] w_{jk\ell}[n] \otimes \hat{s}_{mk}[n] \\
&= \sum_{k=1}^{K} \sum_{j=1}^{M} \sum_{\ell=0}^{N-1} \sum_{h=0}^{H-1} \hat{s}_{mk}[n] x_j[n-h-\ell] w_{jkl}[n-h] \\
&\simeq a_m[n] \ .
\end{aligned}
$$

where the final approximation holds under the assumption of quasi-static weights of the adaptive filter: $w_{jk\ell}[n-h] \simeq w_{jk\ell}[n]$.[7] From (4.4), by adding the reference signal $x_m[n]$ to $a_m[n]$, an estimate of the error signal is obtained

$$b_m[n] = a_m[n] + x_m[n] \ .$$

Similarly to standard FxLMS, the input of the LMS block is, in this case, the estimate of the error signal $b_m[n]$. Therefore, the tap weight equation (3.9), derived in Section 3.1, still holds for the modified FxLMS with $e_m[n]$ replaced by $b_m[n]$.

4.2 Fixed FB Systems

Unlike classic adaptive solutions discussed in the previous sections, a fixed FB controller for ANC purposes is here presented. The idea is to design a filter with a fixed frequency behavior, which does not track instantaneous variations in the audio signals. Such a controller may be of potential interest, since it can be used to cancel in a given frequency range. Note that further mathematical derivation of the FB system and the Wiener solution as an alternative fixed FB ANC system are discussed in Appendices C and D, respectively.

The rest of this section is organized as follows. In Section 4.2.1, a brief review on control theory, needed for the derivation of the fixed controller, is given. In Section 4.2.2, the considered fixed FB controller is described.

[7] For a linear time varying filter cascaded with a linear time invariant one, the commutative property does not hold. In Fig. 4.3, the order of filtering to obtain $y'_m[n]$ and $a_m[n]$ from $x_j[n]$ is exchanged.

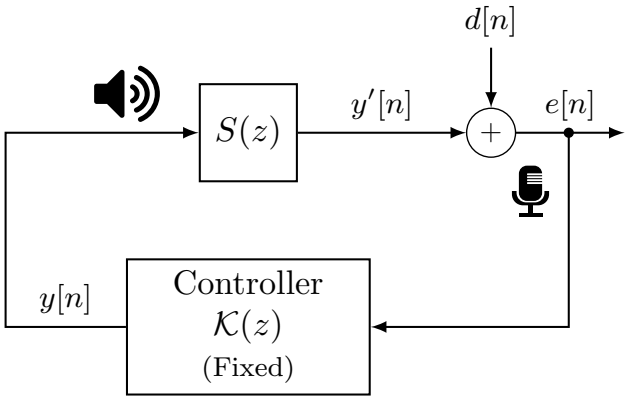


Figure 4.4: Block diagram of a fixed FB ANC system.

4.2.1 Review of Control Theory

In Fig. 4.4, a general fixed FB control system is depicted. In particular, the filter $S(z)$ represents the secondary path (also known as *plant* in control theory) and the controller $\mathcal{K}(z)$ represents a fixed filter. From Fig. 4.4, the error signal is calculated as

$$\begin{aligned} e[n] &= d[n] + y'[n] \\ &= d[n] + y[n] \otimes s[n] \\ &= d[n] + e[n] \otimes h[n] \otimes s[n] \end{aligned} \tag{4.5}$$

where $h[n]$ is the impulse response of the fixed filter $\mathcal{K}(z)$. By performing the $\mathcal{Z}-$transform of (4.5), one gets

$$E(z) = D(z) + E(z)\mathcal{K}(z)S(z) \ .$$

Rearranging terms, one has,

$$E(z) = T(z)D(z) \tag{4.6}$$

where $T(z)$ identifies the transfer function[8] between the input $d[n]$ and the output signal $e[n]$ which can be expressed as

$$T(z) = \frac{1}{1 - \mathcal{K}(z)S(z)} \ .$$

[8]In control theory and systems, $T(z)$ is usually referred to as *sensitivity function*, since it gives us information on how much the disturbances are influenced by the feedback.

For the sake of simplicity and following the literature, the mathematical derivations are formulated in continuous time in terms of Fourier transform $\mathcal{F}\{\cdot\}$. Extending the Fourier domain to the $\mathcal{Z}$-domain is then straightforward. Therefore, (4.6) becomes

$$E(j\omega) = T(j\omega)D(j\omega) \tag{4.7}$$

being $T(j\omega)$ the frequency domain version in continuous time of the transfer function $T(z)$ defined by analogy as

$$T(j\omega) = \frac{1}{1 - \mathcal{K}(j\omega)S(j\omega)} \ .$$

In the literature, several and different approaches were proposed for designing the controller $\mathcal{K}(z)$, such as the so-called *mixed-sensitivity* $\mathcal{H}_\infty$ control method [30], based in the Hardy-space, which is used for ANC purposes [31–33]. The idea behind $\mathcal{H}_\infty$ is to express the control problem as an optimization one. Once the mathematical optimization is solved, a synthesized controller with guaranteed performance is obtained. The objective of $\mathcal{H}_\infty$ control design is to find a controller $\mathcal{K}(z)$ such that the $\mathcal{H}_\infty$ norm of the transfer function $T(j\omega)$ is minimized. The $\mathcal{H}_\infty$ norm is calculated as

$$\|T(j\omega)\|_\infty = \sup_\omega \sigma_{\max}\left(T(j\omega)\right)$$

where $\sigma_{\max}$ is the maximum singular value of the transfer function $T(j\omega)$.

Similarly to the $\mathcal{H}_\infty$ control method, $\mathcal{H}_2$ and a combination of the aforementioned control methods, the so-called $\mathcal{H}_2/\mathcal{H}_\infty$ was proposed and analyzed in [34]. In case of adopting the $\mathcal{H}_2$ control method, the $\mathcal{H}_2$ norm of the transfer function $T(j\omega)$ is defined as

$$\|T(j\omega)\|_2^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{tr}\left\{T^*(j\omega)T(j\omega)\right\}$$

where $\mathrm{tr}\{\cdot\}$ is the trace operator and $*$ denotes the conjugate operator.

4.2.2 Considered Fixed FB Controller

A simplified and intuitive solution of fixed FB controller was considered for our MATLAB-based simulator. More precisely, a slightly different version of the controller proposed

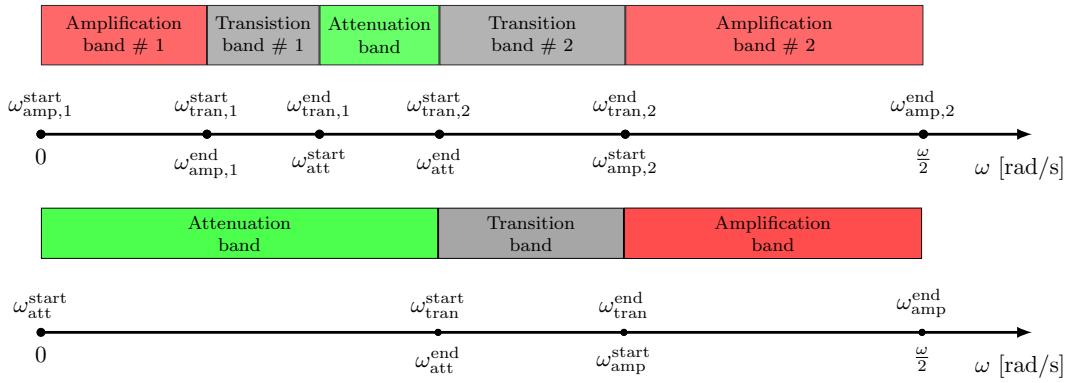


Figure 4.5: Frequency band division for a band-pass filter (top) and for a low-pass filter (bottom).

in [35] was implemented and it is here discussed.

The so-called *waterbed effect* says that a noise attenuation over a certain frequency range causes a noise amplification outside that considered bandwidth [36]. In order to cope with this effect, the idea behind [35] is to design a filter $\mathcal{K}(j\omega)$, such that the open-loop transfer function $\mathcal{K}(j\omega)S(j\omega)$ behaves like a Band-Pass Filter (BPF) or Low-Pass Filter (LPF) within the attenuation frequency range and is such that the closed-loop transfer function is flat outside that range.

The frequency bandwidth can be thus divided into attenuation band Ω_{Att}, around the energy peaks of the disturbance signal, and amplification bands Ω_{Amp}, otherwise. For stability purposes, in order to avoid discontinuities in the open-loop transfer function at the border of attenuation and amplification bands, a proper interpolation function $\mathcal{R}(j\omega)$ in the so-called transition bands Ω_{Tran} may be inserted. The frequency bandwidth can be thus divided into attenuation bands, transition bands and amplification bands, as shown in Fig. 4.5.

The overall controller design procedure can be summarized as follows.

1. Define a initial filter $H_0(j\omega)$ either LPF or BPF.

2. Design an interim control filter $F(j\omega)$ in order to flatten the closed-loop transfer function in the amplification bands.

3. Design arches of raised cosine $\mathcal{R}(j\omega)S(j\omega)$ in the transition bands.

4. Design the resulting control filter $\mathcal{K}(j\omega)$ and convert it to discrete-time as $\mathcal{K}(z)$.

The objective of ANC systems is to reduce, in magnitude, the output signal $e[n]$. From (4.7), in order to have noise reduction, one needs $|T(j\omega)| < 1$. On the other hand, if $|T(j\omega)| > 1$, the noise is amplified. Following the procedure of fixed controller design detailed above, a filter $H_0(j\omega)$, is defined as

$$H_0(j\omega) = \begin{cases} \frac{j\omega A_p \frac{w_p}{Q_p}}{(j\omega)^2 + j\omega \frac{\omega_p}{Q_p} + \omega_p^2} & \text{if BPF} \\ \frac{A_p}{j\omega \frac{1}{\omega_0} + 1} & \text{if LPF} \end{cases}$$

where, for a BPF, $A_p = H_0(j\omega_p)$ is the magnitude at the center frequency, $w_p = 2\pi f_p$ is the peak cancellation angular frequency, f_p the peak cancellation frequency, $Q_p = \frac{f_p}{B_w}$ the quality factor and B_w the BPF bandwidth. For a LPF, $A_p = H_0(0)$ is set. For compactness, in the following we use the symbol ω_p to denote 0 for a LPF.

It is possible to calculate a-priori the maximum noise attenuation achievable over the noise attenuation band Ω_{att} going from $\omega_{\text{att}}^{\text{start}}$ to $\omega_{\text{att}}^{\text{end}}$, i.e., $\Omega_{\text{att}} = \left[\omega_{\text{att}}^{\text{start}}, \omega_{\text{att}}^{\text{end}}\right]$. Hence, the Noise Reduction (NR) $\eta(\omega)$ is defined as

$$\begin{aligned} \eta(\omega) &= -20 \log_{10} |T(j\omega)| \\ &= -20 \log_{10} \left| \frac{1}{1 - H(j\omega)S(j\omega)} \right| \\ &= 20 \log_{10} |1 - H(j\omega)S(j\omega)| \qquad \omega \in \Omega_{\text{att}} \ . \end{aligned} \tag{4.8}$$

The larger noise reduction $\eta(\omega)$, i.e., the smaller $T(j\omega)$, the greater the noise cancellation. By evaluating (4.8) at w_p one can find the value of A_p given a desired value of $\eta(\omega_p)$. This yields

$$\eta(\omega_p) = 20 \log_{10} |1 - A_p S(j\omega_p)| \ . \tag{4.9}$$

Note that (4.9), for BPF, has two roots and the correct solution in terms of A_p is selected such that the following condition is satisfied

$$|S(j\omega)H_0(j\omega)| > 1 \qquad \omega \in \Omega_{\text{att}}$$

where $H_0(j\omega)$ is the initial filter computed with the corresponding A_p. As previously

mentioned, a flat magnitude response is set over the amplification bands. For the whole set of bands

$$\Omega_{\text{amp}} = \begin{cases} \left[\omega_{\text{amp},1}^{\text{start}}, \omega_{\text{amp},1}^{\text{end}}\right] \cup \left[\omega_{\text{amp},2}^{\text{start}}, \omega_{\text{amp},1}^{\text{end}}\right] & \text{for BPF} \\ \left[\omega_{\text{amp}}^{\text{start}}, \omega_{\text{amp}}^{\text{end}}\right] & \text{for LPF} \end{cases}$$

the corresponding average noise amplification level ν, in dB, can be calculated a

$$\nu = \frac{1}{B_{\text{amp}}} \int_{\omega \in \Omega_{\text{amp}}} 20 \log_{10} \frac{1}{|1 - H_0(j\omega)S(j\omega)|} \, d\omega \ ,$$

where B_{amp} is the overall size of the noise amplification band Ω_{amp}. Using the Bode's sensitivity integral [37] defined as

$$\int_0^{+\infty} 20 \log_{10} |T(j\omega)| \ d\omega = 0$$

the following expression can be derived

$$\begin{aligned} \nu &= \frac{1}{B_{\text{amp}}} \int_{\omega \in \Omega_{\text{amp}}} 20 \log_{10} \frac{1}{|1 - H_0(j\omega)S(j\omega)|} \, d\omega \\ &= \frac{1}{B_{\text{amp}}} \int_{\omega \in \Omega_{\text{amp}}} 20 \log_{10} |T(j\omega)| \ d\omega \\ &= \frac{1}{B_{\text{amp}}} \left[\underbrace{\int_0^{\infty} 20 \log_{10} |T(j\omega)| \ d\omega}_{=0} - \int_{\omega \in \Omega_{\text{att}}} 20 \log_{10} |T(j\omega)| \ d\omega \right] \\ &= -\frac{1}{B_{\text{amp}}} \int_{\omega \in \Omega_{\text{att}}} 20 \log_{10} |T(j\omega)| \ d\omega \\ &= \frac{1}{B_{\text{amp}}} \int_{\omega \in \Omega_{\text{att}}} 20 \log_{10} |1 - H_0(j\omega)S(j\omega)| \ d\omega \ . \end{aligned}$$

Note that if $|1 - H_0(j\omega)S(j\omega)|$ is constant for $\omega \in \Omega_{\text{amp}}$, then it equals $10^{-\nu/20}$.

The magnitude of the transfer function $T(j\omega)$ should be

$$\begin{aligned} |T(j\omega)| &= \frac{1}{|1 - F(j\omega)S(j\omega)|} \\ &= |1 - F(j\omega)S(j\omega)|^{-1} \\ &= 10^{\frac{\nu}{20}} \geq 1 \qquad\qquad \omega \in \Omega_{\text{amp}} \ . \end{aligned}$$

As a simplification (see [35]), the filter $F(j\omega)$, in the amplification bands Ω_{amp}, is defined as

$$F(j\omega) = \frac{1 - 10^{-\frac{\nu}{20}}}{|S(j\omega)|} e^{j\phi_F(\omega)} \quad \omega \in \Omega_{\text{amp}}$$

where $\phi_F(j\omega)$ is the phase of the filter, defined starting from

$$\phi_F(\omega) = \angle H_0(j\omega) \qquad \omega \in \Omega_{\text{amp}}$$

by the following procedure. From control theory we know that for stability reasons and in order to flatten the noise amplification in the corresponding bands, the phase limit of the open-loop frequency response must satisfy

$$\ell\pi - \theta < \arg\left[S(j\omega)F(j\omega)\right] < \ell\pi + \theta \qquad \omega \in \Omega_{\text{amp}} \quad \ell = 0, 2, 4, ... \tag{4.10}$$

where θ is evaluated from the open-loop Nyquist plot and is equal to

$$\theta = \arccos\left(\frac{1 - 10^{-\frac{\nu}{20}}}{2}\right) \ .$$

If the constraint in (4.10) is not satisfied, a phase lead/lag can be introduced in the phase of $F(j\omega)$.

In order to get a smoother response between amplification and attenuation bands, arches of raised cosine, in the magnitude of the filter, could be inserted within the transition bands, i.e., $\Omega_{\text{tran},i} = \left[\omega_{\text{tran},i}^{\text{start}}, \omega_{\text{tran},i}^{\text{end}}\right]$ for $i = 1$ or $i = 1, 2$ for LPF and BPF, respectively. Therefore, similarly to the amplification bands, the filter is defined as

$$\mathcal{R}_i(j\omega) = |\mathcal{R}_i(j\omega)|\, e^{j\phi_{\mathcal{R}_i}(\omega)} \quad \omega \in \Omega_{\text{tran},i}$$

where the magnitude

$$|\mathcal{R}_i(j\omega)| = b_i + \frac{1}{2}(a_i - b_i)\left\{1 + \cos\left[\pi\left(\frac{\omega_i - \omega_{a_i}}{\omega_{b_i} - \omega_{a_i}}\right)\right]\right\} \quad \omega \in \Omega_{\text{tran},i} \tag{4.11}$$

and the phase $\phi_{\mathcal{R}_i}(\omega)$ is set to be linear. In particular, for stability, a junction between the amplification band(s) and the attenuation band is properly done, both for the magnitude and the phase.

If a BPF is adopted, in (4.11), one has

$$\begin{aligned}
a_1 &= \left|F_1(j\omega_{\text{amp},1}^{\text{end}})\right| && \text{and} && \omega_{a_1} = \omega_{\text{amp},1}^{\text{end}} \\
b_1 &= |H_0(j\omega_{\text{att}}^{\text{start}})| && \text{and} && \omega_{b_1} = \omega_{\text{att}}^{\text{start}} \\
a_2 &= \left|H_0(j\omega_{\text{att}}^{\text{end}})\right| && \text{and} && \omega_{a_2} = \omega_{\text{att}}^{\text{end}} \\
b_2 &= \left|F_2(j\omega_{\text{amp},2}^{\text{start}})\right| && \text{and} && \omega_{b_2} = \omega_{\text{amp},2}^{\text{start}} \ .
\end{aligned}$$

Similar considerations hold for the LPF.

Finally, the interim filter can be written as

$$\mathcal{K}(j\omega) = \begin{cases} F(j\omega) & 0 < \omega < \omega_{\text{amp},1}^{\text{end}} \quad \text{and} \quad \omega_{\text{amp},2}^{\text{start}} < \omega < \omega_{\text{amp},2}^{\text{end}} \\ \mathcal{R}_1(j\omega) & \omega_{\text{tran},1}^{\text{start}} < \omega < \omega_{\text{end}}^{\text{tran},1} \\ H_0(j\omega) & \omega_{\text{att}}^{\text{start}} < \omega < \omega_{\text{end}}^{\text{att}} \\ \mathcal{R}_2(j\omega) & \omega_{\text{tran},2}^{\text{start}} < \omega < \omega_{\text{tran},2}^{\text{end}} \end{cases} .$$

Following the fourth point of the procedure mentioned above, a design of the filter is thus performed in order to get a controller $\mathcal{K}(j\omega)$. The MATLAB function `fdesign.arbmagnphase` was used for this task. In particular,

- `fdesign.arbmagnphase(n,b,f,h)` for Finite Impulse Response (FIR) filter
- `fdesign.arbmagnphase(na,nb,f,h)` for Infinite Impulse Response (IIR) filter

where `n` identifies the filter order, `b` the number of bands, `nb` and `na` numerator and denominator orders respectively, `f` the frequency vector, and finally `h` identifies the complex frequency response values.

Once the filter $\mathcal{K}(j\omega)$ is evaluated, it can be converted into $\mathcal{K}(z)$ in the fixed controller shown in Fig. 4.4.

Chapter 5

Types of Input Signals

In this chapter, the input signals considered for our simulations are presented. More precisely, as a preliminary experiment, synthetic tonal input signals, shown in Section 5.1, are employed. Subsequently, in Section 5.2, a set of experimental microphone acquisitions is presented.

Due to the absence of reference signals, a FB ANC system tries to cancel the perceived audio signal in the overall spectrum, exploiting the feedback of the error signal for reference signal synthesis purposes. This means that the system performance is strongly dependent on the nature of the input signal. For instance, a reconstruction of a fast time-varying input signal implies negligible loop delay in the system. As a consequence, FB ANC systems are better suited for tonal signal cancellation (e.g., the noise caused by the engine speed). In the following subsections, a brief description on the synthetic and experimental input signals for the considered simulations is given.

5.1 Synthetic Tonal

As previously mentioned, FB ANC works efficiently when the input signal is tonal or, at least, with a limited number of narrow-band components. Therefore, two types of input signals were synthesized:

- Sum of sinusoids.
- Narrow-band filtered white noise.

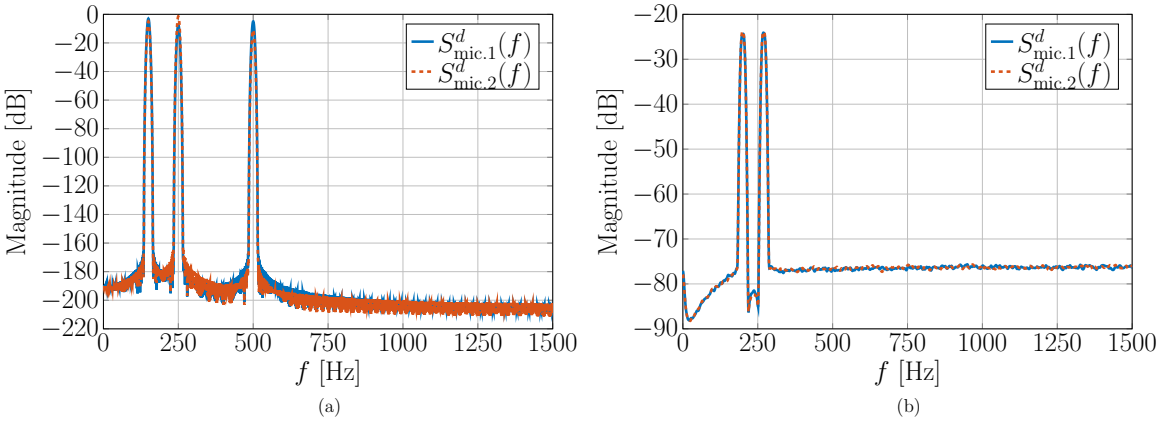


Figure 5.1: SPL spectra of synthesized tonal input signals for error microphone 1 and 2. (a) sum of sinusoids and (b) narrow-band filtered white noise.

In particular, in case of sum the sinusoids, the input signal is

$$d[n] = \sum_{i=1}^{3} A_i \cos\left(\frac{2\pi f_i n}{f_s} + \phi_i\right)$$

with arbitrary amplitudes $A_i = 1, 0.6, 0.8$, frequencies $f_i = 150, 250, 500$ Hz and phases $\phi_i = 0, 1, 2$ rad for $i = 1, 2, 3$ respectively. For both types, the sampling frequency is $f_s = 3$ kHz. The considered pass-band center frequencies for filtered white noise input signal are 200 and 270 Hz with a bandwidth of few Hz (10–20 Hz). The SPL spectra of the considered tonal input signals are depicted in Fig. 5.1 for both error microphones. Three independent sinusoids are included, and two different very narrow-band filters are employed for filtering the white noise.

5.2 Experimentally Acquired Signals

For our simulations, two types of experimentally acquired input signals, obtained in a sedan and in a hybrid compact car (class C), are discussed in the following.

5.2.1 Sedan

The SPL spectra of employed input signal are shown in Fig. 5.2 for the error microphones 1 and 2. These experimental signals, sampled at 3 kHz, come from a measure-

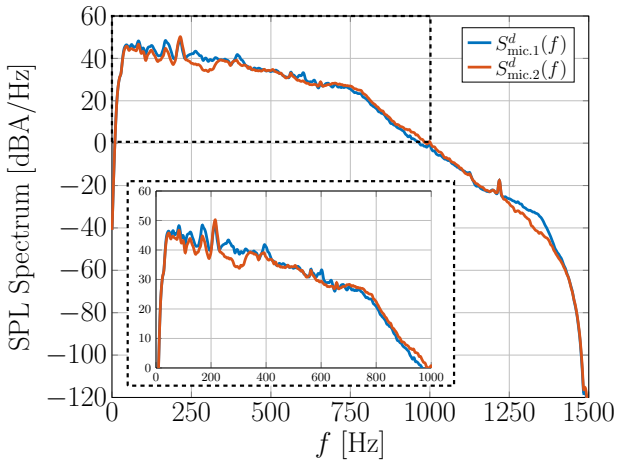


Figure 5.2: SPL spectra of experimental error microphone signals in sedan for error microphones 1 and 2.

ment campaign performed by ASK Industries in a sedan and are the same used and shown in Section 3.2 for FF ANC systems. In particular, a manikin equipped with binaural microphones was placed in the passenger's seat.

5.2.2 Hybrid compact car

In Fig. 5.3, the spectra of the last considered input signal is depicted for the error microphones 1 and 2. In this case, the employed car for this experimental measurement campaign, performed by ASK Industries, is a hybrid compact car.

The setup consists of binaural microphones inserted in a manikin placed at the driver's seat and is used for recording, for about 18 seconds, the noise produced by the hybrid compact car in idle engine conditions with about 800–900 rpm. The used sampling frequencies f_s are 3 kHz and 6 kHz. However, only $f_s = 3$ kHz is considered for our analysis. This signal is representative of a roughly tonal input, where the harmonics are caused by the rotations of the engine. Note that the spectra in Fig. 5.3 are those of roughly periodic signals with a fundamental frequency of about 28 Hz, i.e., that of the idle engine. Due to this periodicity, this signal is potentially interesting to be employed for ANC purposes in FB schemes. Moreover, in order to improve the system performance and cast it in a realistic scenario, filtered versions of these signals, within the band 80–400 Hz, with SPL spectra depicted in Fig. 5.4, are employed and

the relevant system performance analyzed.

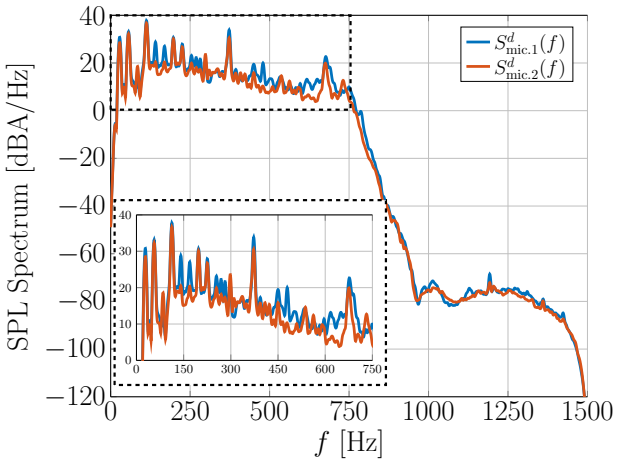


Figure 5.3: SPL spectra of experimental error microphone signals in hybrid compact car.

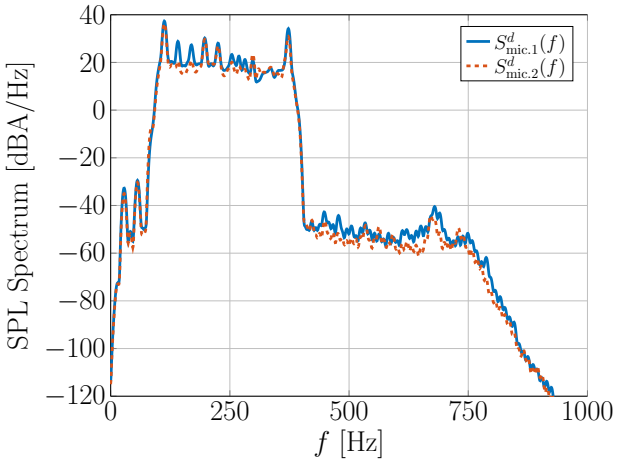


Figure 5.4: SPL spectra of filtered version of the recorded error microphone signals in hybrid compact car.

Chapter 6

Numerical Results

In this chapter, numerical results for the FB ANC schemes considered in Chapter 4 are presented and discussed. The chapter is structured as follows. Simulation scenarios and performance indicators are presented in Section 6.1. Numerical results are organized based on the secondary paths employed in the simulations. More precisely, in Section 6.2 the secondary paths are assumed to be pure delays. In Section 6.3, the secondary paths are the same as those presented in Chapter 2, i.e., those from the experimental measurements performed by ASK in the sedan. Finally, a new experimental setup is considered in Section 6.4, composed by two loudspeakers and two microphones positioned in the headrest of an office chair. In all the cases with adaptive approach, numerical results are obtained by using a value of the step-size parameter which approximately maximizes the overall system performance.

6.1 Scenarios

The FB ANC systems described in Chapter 4 are simulated, namely, SISO and MIMO scheme. Further results and for SISO and MIMO systems is provided and the analysis of the so-called CICO scheme, as a valid alternative to the MIMO system, is presented in Appendices E.1, E.2 and E.3, respectively. Moreover, as explained in Chapter 4, it is possible to distinguish the systems based on the type of controller $\mathcal{K}(z)$ that is implemented, i.e., adaptive or fixed approaches. The SISO FB ANC system, unlike the MIMO one, is analyzed for both the adaptive and fixed approaches. The considered

Table 6.1: Considered scenarios for FB ANC systems.

Scheme \ $S(z)$	DELAY	SEDAN	HEADREST CHAIR
SISO	◔ ■ ◇	◓ ◇	◔ ■ ◇
MIMO	◔ ● ◇	◓ ◇	◓ ◇

◔ Sedan input signal — ◔ hybrid compact car input signal — ◓ Sedan and filtered hybrid compact car input signal — ◓ hybrid compact car and filtered hybrid compact car input signal — ● Filtered white noise input signal — ■ Fixed ANC Approach — ◇ Adaptive ANC Approach.

scenarios and relative input signals are summarized in Table 6.1.

Similarly to FF ANC systems in Section 2.2, system performance is assessed in terms of SPL spectra measured in dBA relative to $p_0 = 20$ μPa with a RBW of 6 Hz. Moreover, in order to analyze the ANC system performance under a different viewpoint, sliding window SPL values of disturbance and error signals are also measured. For a pressure signal $p[n]$, the sliding window SPL value (in dBA) is defined as

$$\mathrm{L}_p[n] = 10 \log_{10} \frac{1}{L p_0^2} \sum_{\ell=1}^{L} p^2[n-\ell] \tag{6.1}$$

where L is the considered window length.

6.2 Pure Delay as a Secondary Path

In this case, we consider the schemes of Chapter 4 with $S(z) = z^{-n_0}$, i.e., a pure delay of n_0 samples. By considering the fixed approach, (4.9) becomes

$$\eta = 20 \log_{10} \left| 1 - A_p e^{-j\omega\tau} \right| \tag{6.2}$$

where $e^{-j\omega\tau}$ is the Fourier transform of a delay $\tau = n_0/f_s$, being n_0 the delay in number of samples.[9] Therefore from (6.2), it is possible to find A_p for a given noise attenuation

[9]For a single sample delay, one has $n_0 = 1$. Since the sampling frequency is $f_s = 3$ kHz, $\tau = 0.33$ ms results.

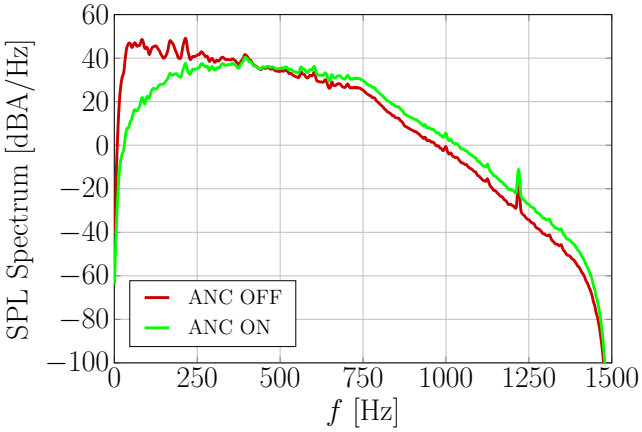


Figure 6.1: SPL spectra for SISO adaptive FB ANC system for OFF and ON cases and left error microphone with sedan input signal.

goal η from

$$\underbrace{10^{\frac{\eta}{20}}}_{=\eta_{\text{lin.}}} = \left|1 - A_p e^{-j\omega\tau}\right|$$

$$\eta_{\text{lin.}} = \left|1 - A_p \cos(\omega\tau) + j \sin(\omega\tau)\right|$$

$$\eta_{\text{lin.}}^2 = A_p^2 - 2A_p \cos(\omega\tau) + 1$$

and solving with respect to A_p.

In Fig. 6.1, the SPL spectra are shown as a function of frequency for FB ANC system ON and OFF. The results are relative to the case of the adaptive approach for a SISO system with sedan input signal of error microphone 1. We can immediately notice that a significant noise cancellation occurs at low frequencies. In the high frequency regime (e.g., 750–1500 Hz), the green curve is above the red one denoting a degradation of performance. This is mostly due to the narrow-band behavior of the FxLMS algorithm that always works around the frequency range in which the peak energy occurs. However, note that at high frequencies, the SPL spectrum values are significantly lower than at low frequencies. For this reason, the noise amplification, at high frequency can be neglected.

In Fig. 6.2, the SPL spectra are shown as a function of frequency for a MIMO

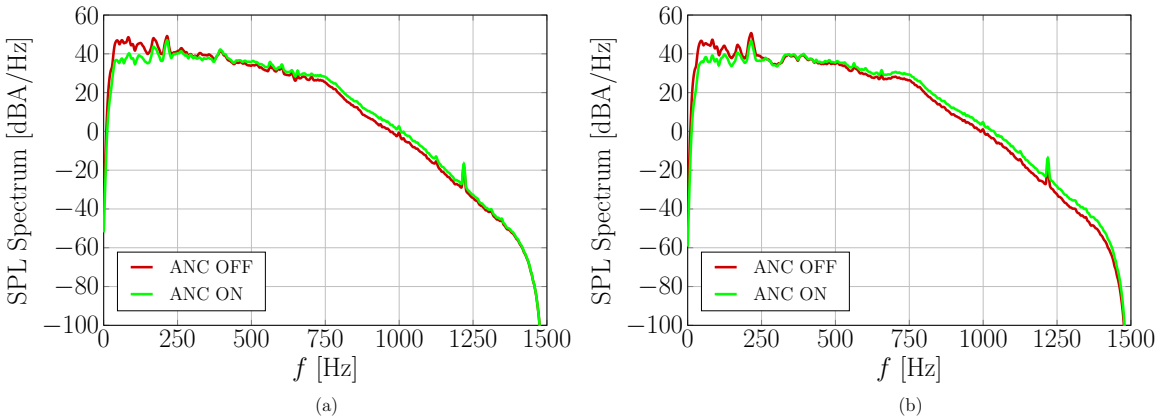


Figure 6.2: SPL spectra of the MIMO adaptive FB ANC system. (a) Left and (b) right error microphone with the sedan input signal.

FB ANC system. For this simulation, $S_{11}(z) = S_{11}(z) = z^{-1}$ and $S_{12}(z) = S_{21}(z) = 0.5z^{-2}$ to emulate different lengths of direct and cross paths. Even if unrealistic, these secondary paths enable to assess the cancellation performance limit. The left (a) and right (b) error microphones are considered. As expected, the obtained results show considerable performance degradation with respect to the SISO case. In fact, it is possible to observe that the presence of the double channel yields a reduced noise cancellation. However in the low frequency regime, we may note a peak cancellation of about 8 dB.

It can be observed that at $f_s = 3$ kHz we cannot afford a delay greater than 1 sample (results are not shown for conciseness). This implies that, under the assumption of a speed of sound $v_s = 343.8$ m/s measured at 21°C, the length between the positions of error microphone and the loudspeaker must be lower than about 11.46 cm.

Numerical results of SPL spectra as a function of frequency for fixed SISO ANC when the system is OFF and ON, for error microphone 1, are reported in Fig. 6.3. Three types of FIR filters were considered: Narrow Band-Pass Filter (NBPF), in the left side in solid black, LPF in the right side in solid green and Broad Band-Pass Filter (BBPF), in the right side in dashed blue. Both the NBPF and the BBPF work around the peak frequency $f_p = 215$ Hz. It is possible to notice that, similarly to a reference signal in a FF ANC system, the NBPF works very efficiently around f_p and performance is not degraded outside this frequency range. On the right hand-side of Fig. 6.3 a comparison between broad-band and low-pass filter is given. The pass-bands

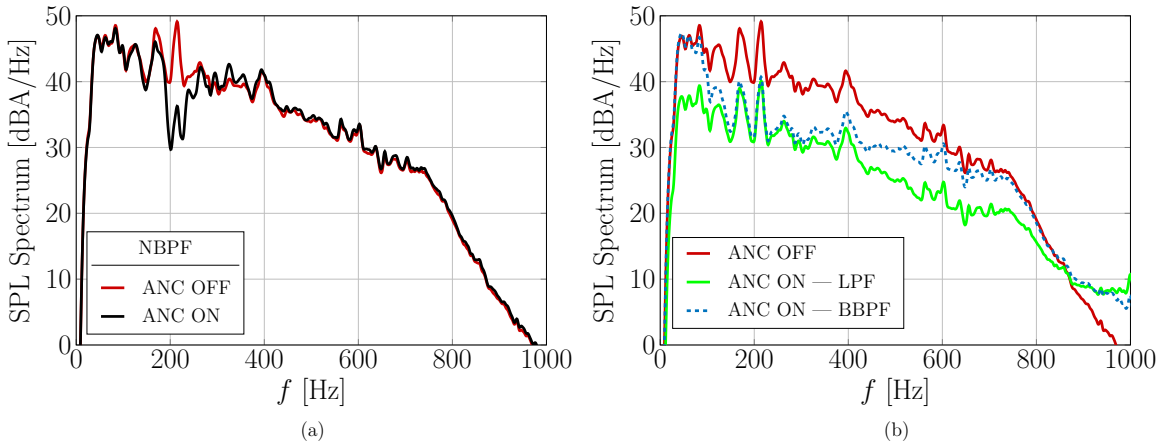


Figure 6.3: SPL spectra with the SISO fixed FB ANC system OFF and ON. Narrow-band FIR filter (a) and comparison between low-pass and broad-band FIR filter (b) for left error microphone with the sedan input signal.

are set from 100 to 800 Hz and from 0 Hz to 800 Hz for BBPF and LPF, respectively. Significant noise cancellation occurs within the pass-band frequency range both for LPF and BPF filter. In particular, LPF is optimal with respect to the BBPF since it guarantees best performance (about 10 dB) without any significant degradation.

6.3 Sedan Secondary Path

In this section, the obtained results related to the sedan secondary path are presented. Both SISO and MIMO FB ANC systems are analyzed for the adaptive approach with sedan and filtered hybrid compact car input signal.

In Fig. 6.4, the SPL spectra of a SISO adaptive FB system is depicted as a function of frequency, for the ANC system OFF and ON. More precisely, the numerical results are related to the case of filtered hybrid compact car input signal for left (a) and right (b) error microphones.

The input signal is very different with respect to the previous results, and the system performance, as expected, is considerably worse, due to the delay introduced by the secondary path of the sedan. The considered secondary paths, in fact, introduce delays of about 18 samples in the direct path and about 22 samples in the cross paths (see Figs. 2.2 and 2.3). For a sampling frequency $f_s = 3$ kHz and a temperature of

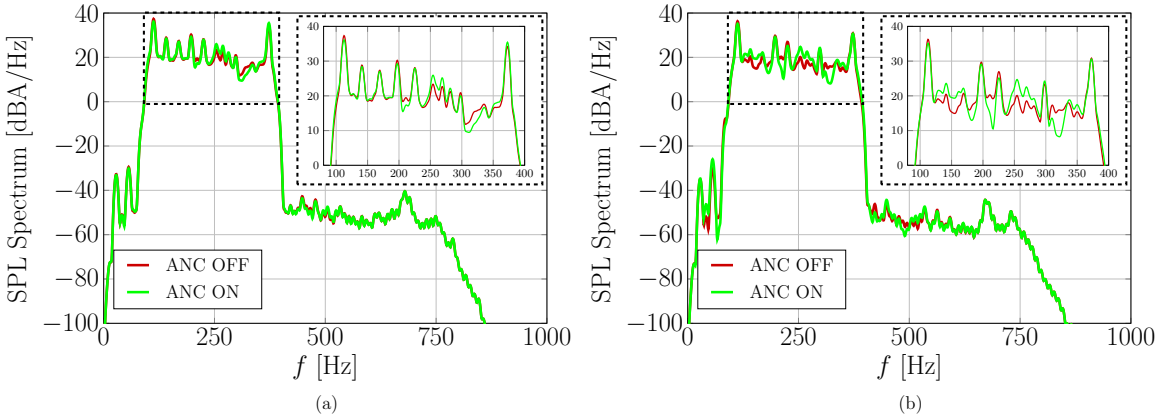


Figure 6.4: SPL spectra with ANC OFF and ON for a SISO adaptive FB ANC system. (a) Left and (b) right error microphone with the hybrid compact car input signal.

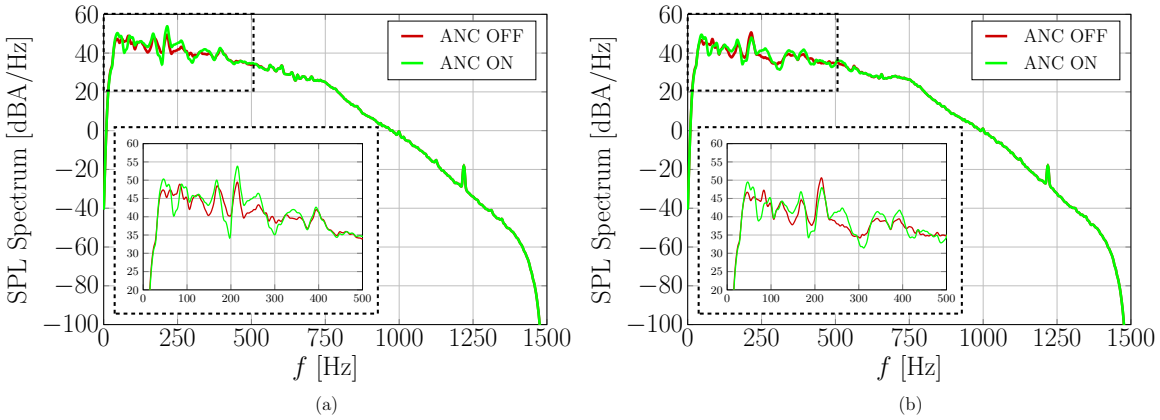


Figure 6.5: SPL spectra with ANC OFF and ANC ON for a MIMO adaptive FB ANC system. (a) Left and (b) right error microphone with the sedan input signal.

21°C, the distance between loudspeakers and error microphones is equal to 206.28 cm and 252.12 cm for the direct and cross paths, respectively. From the previous analysis, we know that the FB system cannot tolerate such a high delay. Moreover, the input signal may be significantly time-varying making the task more complicated.

The SPL spectra of the FB MIMO ANC system for the sedan input signal are shown in Fig. 6.5 for left (a) and right (b) error microphone. It is possible to observe that the system does not exhibit good performance since the green line curve, relative to the ANC system ON, is almost always overlapped with the red one, which refers to the ANC system OFF. Even in this case the input signal is changing too fast over time

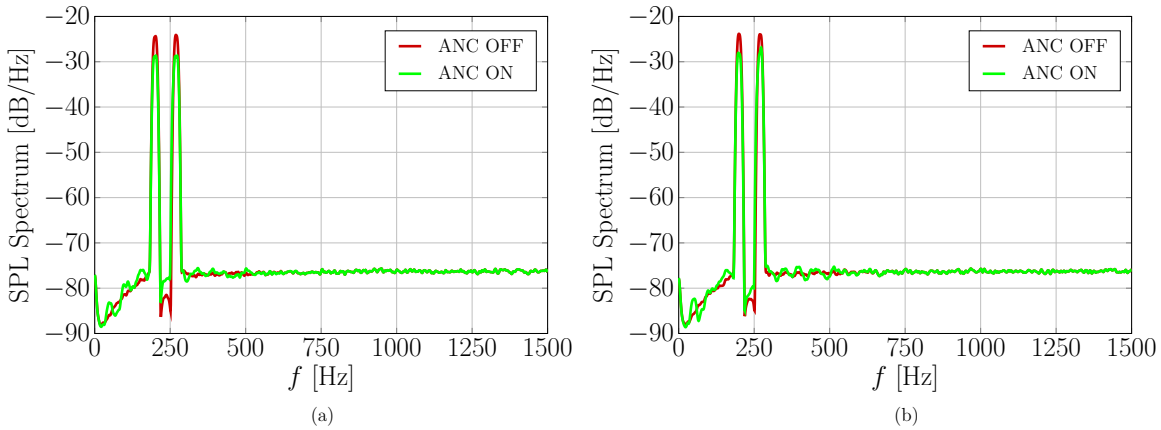


Figure 6.6: SPL spectra with ANC OFF and ANC ON for a MIMO adaptive FB ANC system. (a) Left and (b) right error microphone with the filtered white noise input signal.

and the too large delay introduced by the secondary paths makes reconstruction of the reference signal infeasible. The few noise attenuation instances are always nullified by corresponding amplification ones.

With the aim of testing the FB ANC system in the best possible condition, a tonal input signal is now employed. In particular, the filtered version of white noise, depicted in Fig. 5.1 (b), is used as the input of the MIMO system. The corresponding results, in terms of SPL spectra versus frequency, are reported in Fig. 6.6 for the left (a) and right (b) error microphones . Note that, for this specific simulation, A-weighting was not performed, i.e., the vertical axis is expressed in dB instead of dBA, always relative to $p_0 = 20$ μPa and with RBW of 6 Hz. The peaks around 190 Hz and 270 Hz are well mitigated. This result proves that if the disturbance signal is tonal, the FB ANC system well performs despite the secondary paths introduce significant delays in the loop. Similar considerations can be drawn by considering, as input signal, the sum of sinusoids depicted in Fig. 5.1 (a), and for this reason, performance is not reported.

Figure 6.7 shows the SPL spectra versus frequency for the MIMO adaptive FB ANC system. The employed input signal, in this case, is the filtered version of the hybrid compact car. This signal presents sundry peaks and, for this reason, is suited for a comparison with the standard one (i.e., the input signals shown in Fig. 5.3). It is possible to observe that the noise is slightly mitigated by the system. As observed before, the error microphone 2 experiences a better noise cancellation. The peak noise,

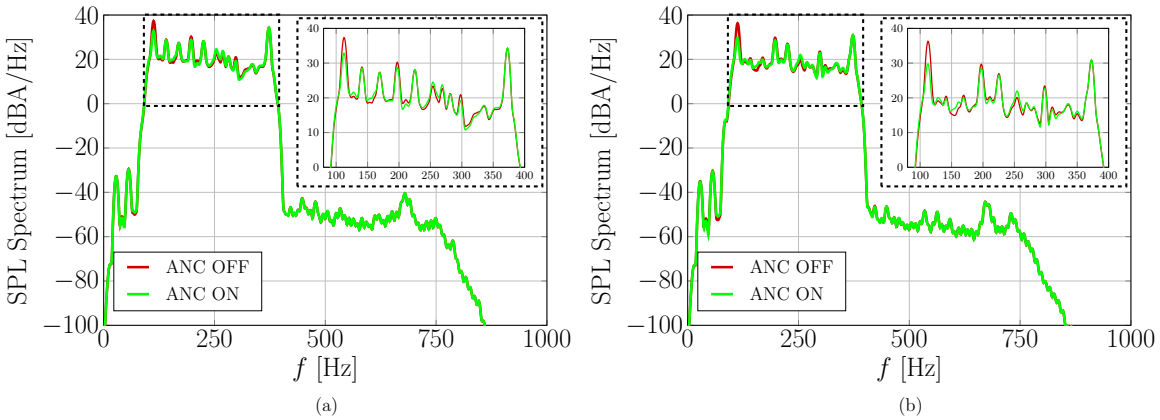


Figure 6.7: SPL spectra with ANC OFF and ANC ON for a MIMO adaptive FB ANC system. (a) Left and (b) right the filtered version of the hybrid compact car input signal.

around 115 Hz, in fact, is properly canceled (about 5 dB of noise cancellation). By considering the left error microphone, the peak cancellation is about 2 dB. However, the system is not performing well since the noise is never canceled except for that region.

For the SISO system, the so-called Filtered Error (FE) approach was also implemented. During our test session, in fact, it was observed that limiting "offline" the bandwidth of the input signal assures some benefits in terms of noise cancellation but is unrealistic from the implementation viewpoint. In this context, with "offline" we mean that the input signal is band limited before running the simulation.

The idea of FE FB is to perform this bandwidth limitation in the control loop. Let us suppose to have a perfect knowledge of the frequency value on which the maximum energy of the noise signal is concentrated and let us call this frequency f_{peak}. It is possible to exploit this information by filtering the error signal at f_{peak} and feeding back this error signal to the LMS algorithm block. In other words, the purpose of FE FB is to obtain a peak cancellation only around f_{peak} and, in this way, to avoid the noise amplification out of the considered band.

The block diagram of the FE FB is depicted in Fig. 6.8, where the aim the NBPF $V(z)$ at the output the error signal is to select a frequency range around f_{peak} as previously described. For consistency, the same filter $V(z)$ is placed at the output of the filter $\widehat{S}(z)$ for the reconstruction of the reference signal. If it is desirable to have

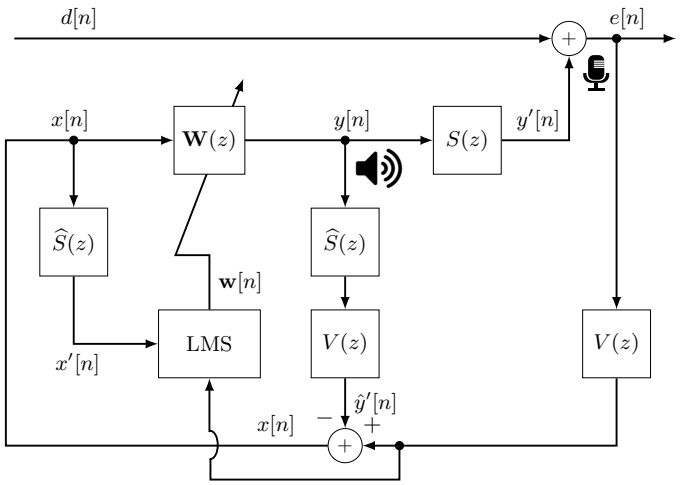


Figure 6.8: Block diagram of FE FB for SISO FB ANC system.

noise cancellation in more than one peak frequency, it is possible to implement a set of narrow-band filters (in parallel with each other) as described later.

The SPL spectra are shown as a function of frequency, for FE FB in Figs. 6.9 and 6.10 for a SISO FB ANC system. In particular, Fig. 6.9 refers to the case of one BPF with a pass band from 92 Hz to 132 Hz, whereas Fig. 6.10 refers to the case of two BPFs from 92 Hz to 132 Hz and 360 Hz to 385 Hz.

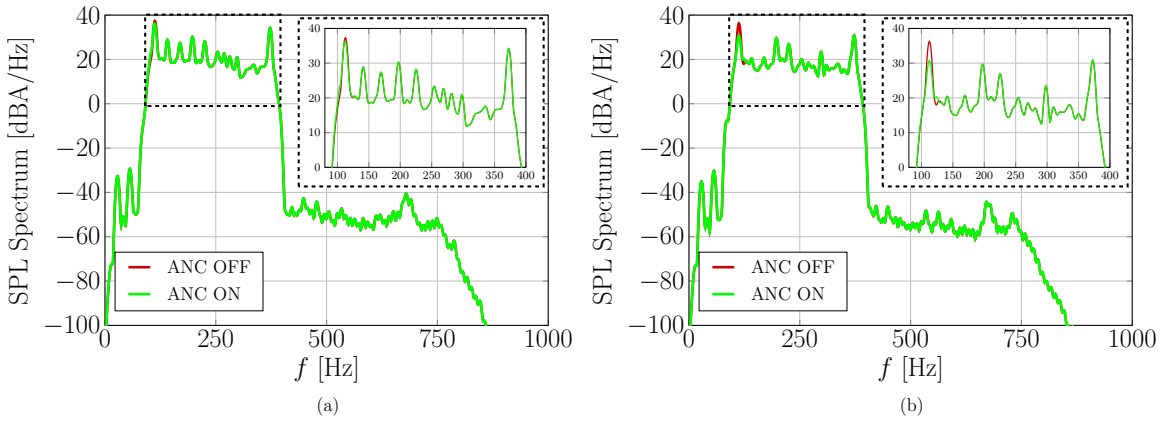


Figure 6.9: SPL spectra with the SISO adaptive FE FB ANC system OFF and ON. (a) Left and (b) right error microphone for one filter and filtered version of hybrid compact car input signal.

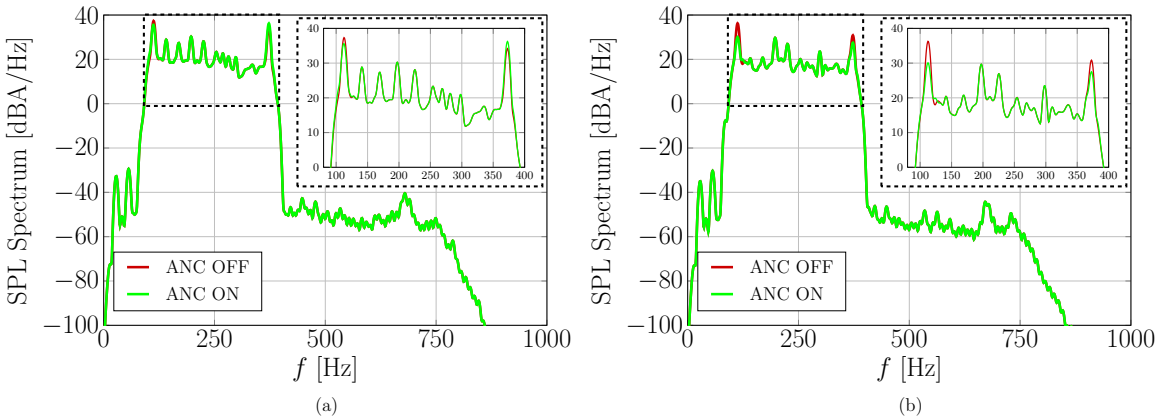


Figure 6.10: SPL spectra with SISO adaptive FE FB ANC system OFF and ON. (a) Left and (b) right error microphones for two filters and filtered version of hybrid compact car input signal.

For both systems, the employed input signal is the filtered hybrid compact car one. The filtered version of the feedback of the error is operating correctly, and, the noise is mitigated in both systems. In particular, it is possible to observe that the right error microphone experiences a better noise cancellation with respect to the left one, as it was also observed in the previous analysis. By considering the right error microphone and the two filters FE FB case, we may note a noise cancellation of about 6 dB and 3 dB for the first and the second peak respectively.

6.4 Headrest Chair Secondary Path

It was observed that the delay introduced by the secondary path has a significant impact on system performance. In order to cope with this issue, it is needed to reduce the distance between error microphones and loudspeakers.

Hence, in collaboration with ASK Industries, an experimental setup composed by an office chair equipped with two loudspeakers and two microphones was realized. In particular, the two loudspeakers were placed in the headrest of the chair and the microphones at the edge of the headrest. The geometrical configuration is depicted in Fig. 6.11, being (a) the front view and (b) the top view.

In Fig. 6.12 the pictures of the prototype are shown. This setup was developed in the listening room of ASK Industries in order to measure the corresponding secondary

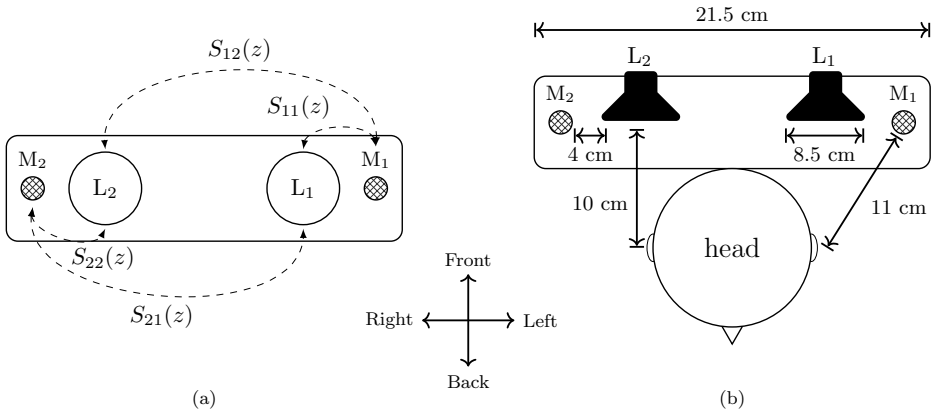


Figure 6.11: Geometrical configuration of the proposed headrest chair: (a) front view and (b) top view.

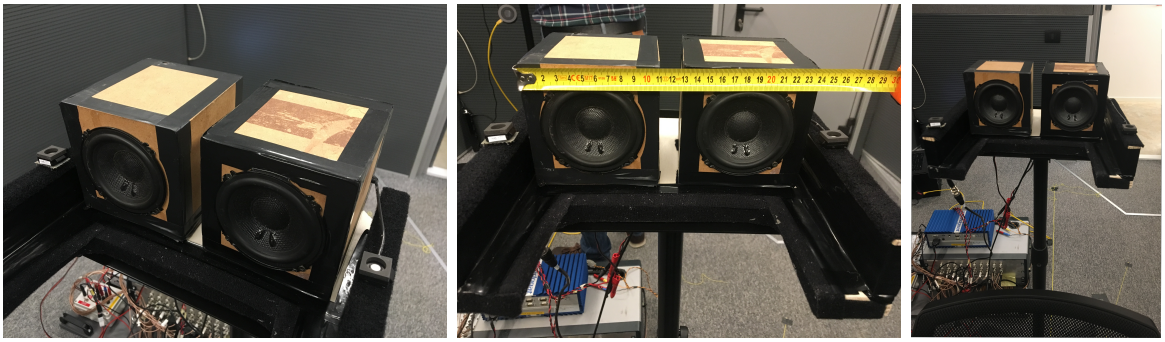


Figure 6.12: Prototype of chair equipped with loudspeakers and microphones within the headrest.

paths $S_{mk}(z)$.

In particular, the estimation of the secondary paths were performed by using an LMS algorithm with white noise as input signal at sampling frequency $f_s = 3$ kHz. The block diagram of the secondary path estimation procedure is depicted in Fig. 6.13.

The audio signal emitted by the loudspeakers is generated at a sampling frequency of 3 kHz, up-sampled to 48 kHz and filtered with a LPF with a cut-off frequency $f_s/2$. At the error microphones, the "received" signal is sent to an anti-aliasing LPF (cut-off frequency set to $f_s/2$) and then down-sampled to $f_s = 3$ kHz. Finally, in order to avoid the Direct Current (DC) component, a high-pass filtering with cut-off frequency at 5 Hz is performed. The error signal, which is given by the subtraction of the convolution of white noise with the secondary path and the output of the high-pass filter, is fed

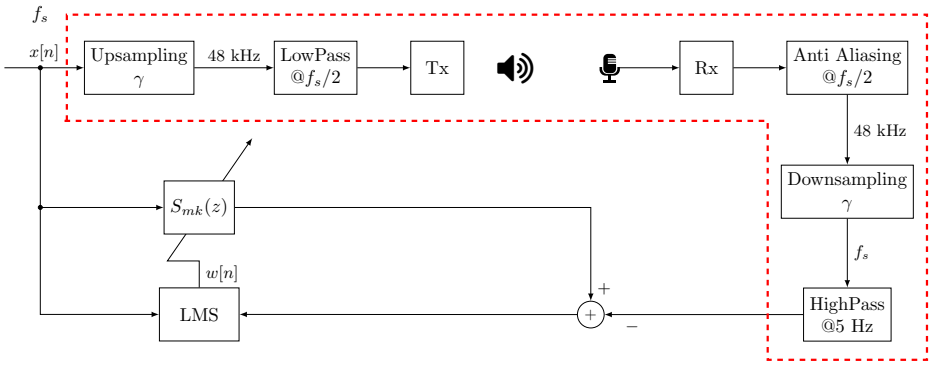


Figure 6.13: Block diagram of the secondary path estimation. The dashed red line identifies the estimated channel.

back the LMS filter in order to update the estimate of the secondary paths.

Two versions of secondary paths estimation were performed based on the type of anti-aliasing filter employed, namely, 8^{th} order IIR LPF with cut-off frequency of 750 Hz and 4^{th} order IIR BPF with a pass-band going from 90 Hz to 390 Hz. For the sake of simplicity, therein they are called 8^{th} ord. and 4^{th} ord. secondary paths and are described in Sections 6.4.1 and 6.4.2, respectively.

6.4.1 8^{th} order Secondary Paths

Magnitude frequency response, impulse response and group delay against frequency graphs of the 8^{th} order headrest chair secondary paths, sampled at $f_s = 3$ kHz, are shown in Fig. 6.14 and Fig. 6.15 for the error microphones 1 and 2, respectively. It is possible to observe that, in the frequency domain, the filters show a band-pass behavior with a cut-off frequency of 750 Hz. Note that the direct channels ($S_{11}(z)$ and $S_{22}(z)$) exhibit a gain greater than that of the cross channel ($S_{12}(z)$ and $S_{21}(z)$). By observing the impulse response, we can note that the filters are represented by 300 taps. In particular, the direct channels introduce a delay of about 4.67 ms, since the response peaks at about the $14-$th sample, and the cross channels a delay of about 15 samples corresponding to 5 ms. Furthermore, note that, within the bandwidth of interest, i.e., 100–750 Hz, the group delays are about 12–15 samples, in agreement with the delays estimated by the impulse responses.

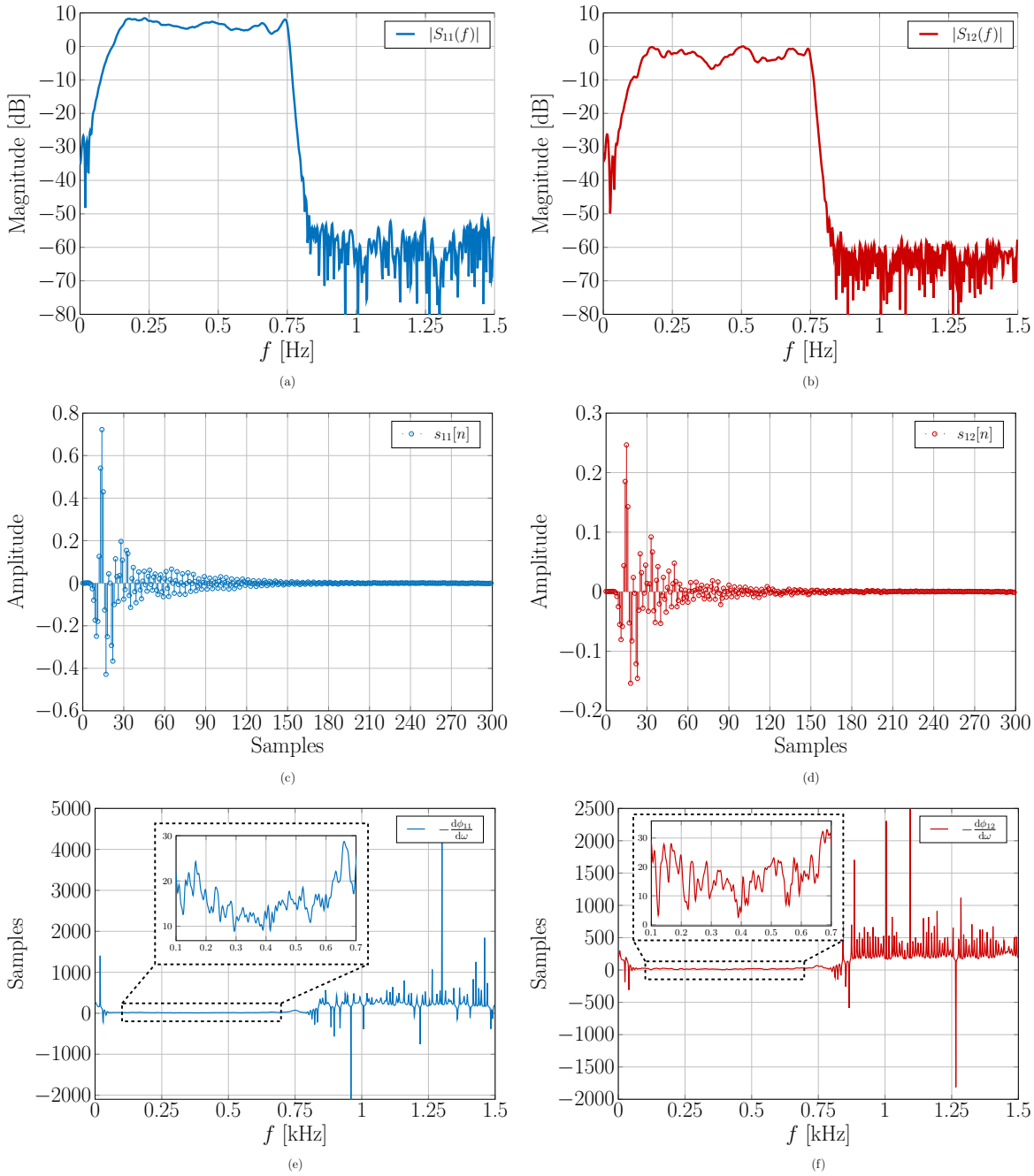


Figure 6.14: Left side: magnitude frequency response (a), impulse response (c) and group delay against frequency (e) of the 8$^{\text{th}}$ order secondary path between loudspeaker 1 and left error microphone. Right side: magnitude frequency response (b), impulse response (d) and group delay against frequency (f) of the 8$^{\text{th}}$ order secondary path between loudspeaker 2 and left error microphone.

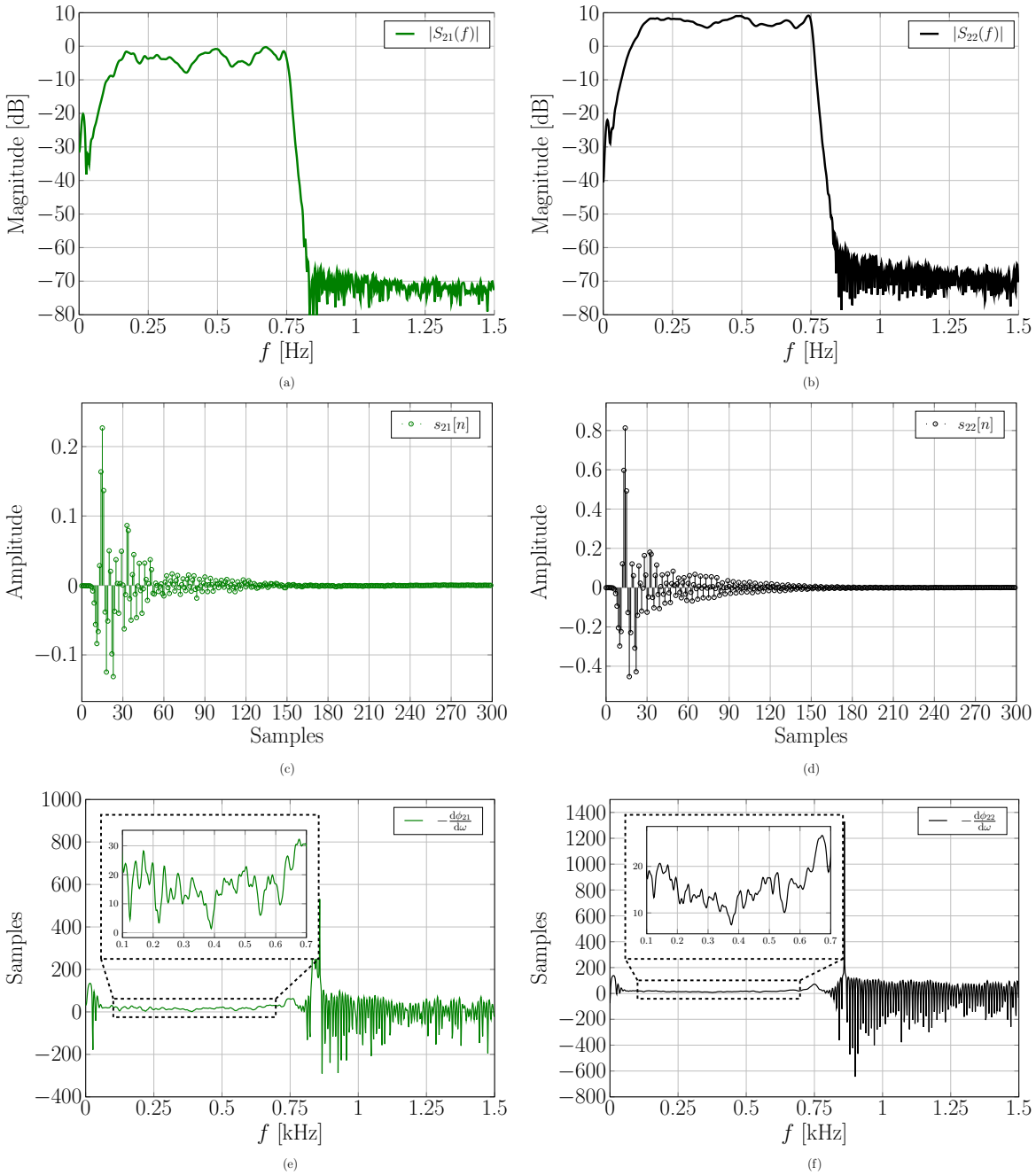


Figure 6.15: Left side: magnitude frequency response (a), impulse response (c) and group delay against frequency (e) of the 8$^{\text{th}}$ order secondary path between loudspeaker 1 and right error microphone. Right side: magnitude frequency response (b), impulse response (d) and group delay against frequency (f) of the 8$^{\text{th}}$ order secondary path between loudspeaker 2 and right error microphone.

6.4.2 4^{th} order Secondary Paths

Magnitude response, impulse response and group delay graphs of the 4^{th} order headrest chair secondary paths, sampled at $f_s = 3$ kHz, are shown in Fig. 6.16 and Fig. 6.17 for left and right error microphones, respectively. The magnitude responses of the secondary paths show a band-pass behavior with a high-frequency cut-off about 700 Hz. Note that the direct acoustic channels, namely $S_{11}(z)$ and $S_{22}(z)$, exhibit a greater gain with respect to the cross acoustic channels, namely $S_{12}(z)$ and $S_{21}(z)$. In particular, by looking at the impulse response, we may observe that the direct channels introduce delays of about 9-10 samples and the cross channels delays of 10–11 samples. At a sampling frequency $f_s = 3$ kHz, these correspond to delays of 3–3.33 ms and 3.33–3.67 ms, respectively. Similar considerations can be drawn by looking at the corresponding group delays.

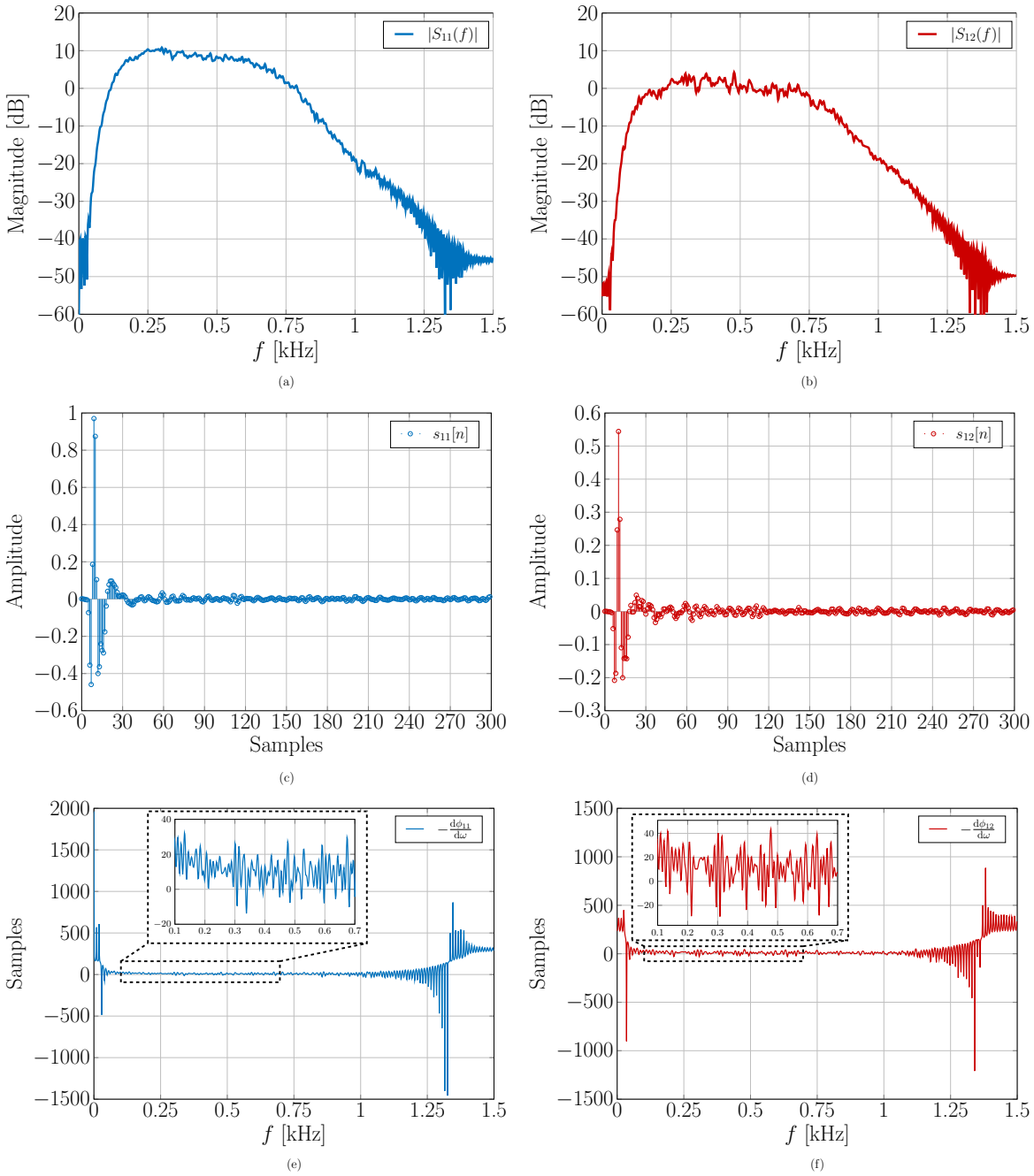


Figure 6.16: Left side: magnitude frequency response (a), impulse response (c) and group delay against frequency (e) of the 4$^{\text{th}}$ order secondary path between loudspeaker 1 and left error microphone. Right side: magnitude frequency response (b), impulse response (d) and group delay against frequency (f) of the 4$^{\text{th}}$ order secondary path between loudspeaker 2 and left error microphone.

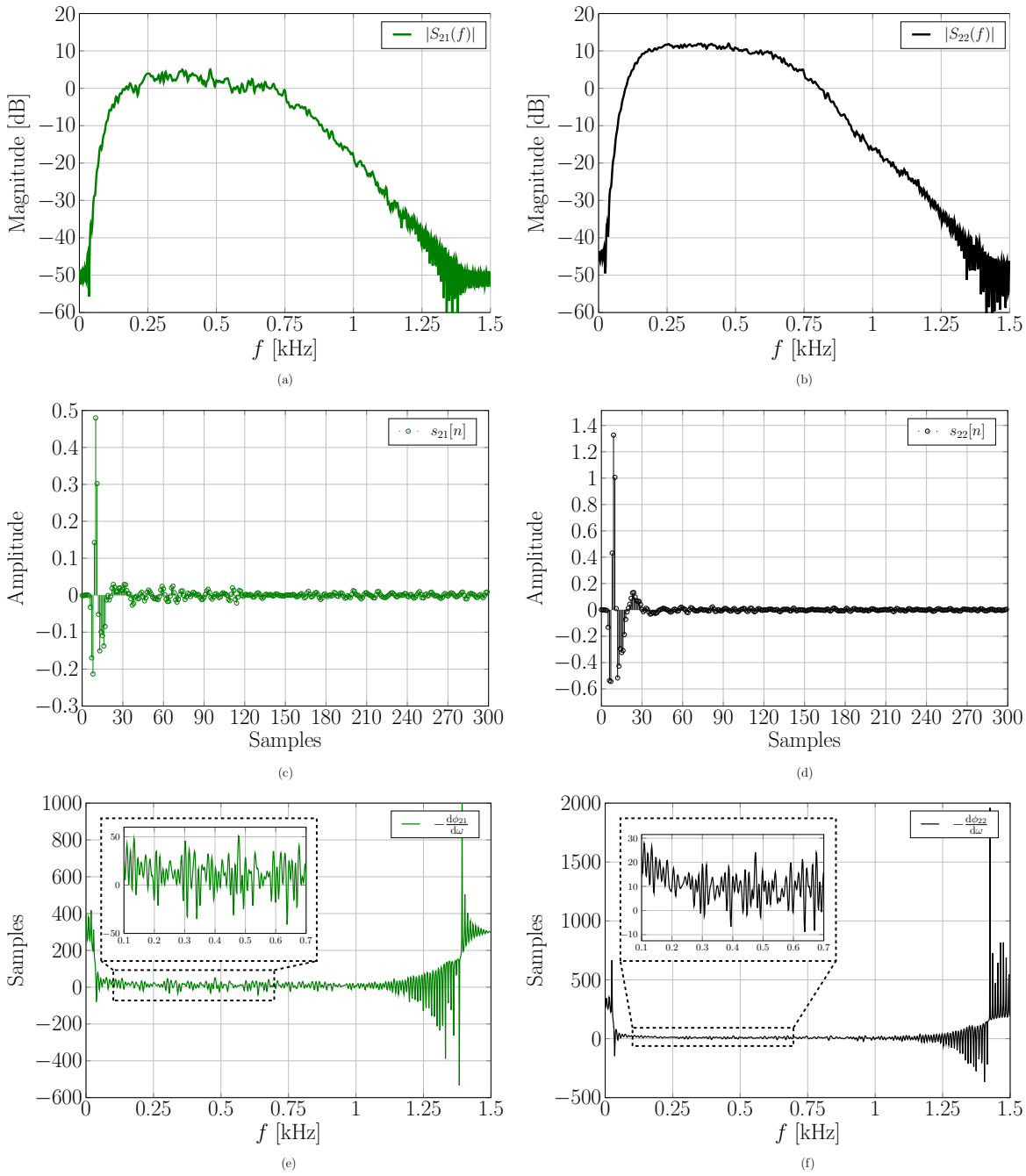


Figure 6.17: Left side: magnitude frequency response (a), impulse response (c) and group delay against frequency (e) of the 4$^{\text{th}}$ order secondary path between loudspeaker 1 and right error microphone. Right side: magnitude frequency response (b), impulse response (d) and group delay against frequency (f) of the 4$^{\text{th}}$ order secondary path between loudspeaker 2 and right error microphone.

Table 6.2: Considered filter parameters.

Left Error Microphone				
	f_p [Hz]	$\eta(\omega_\mathrm{p})$ [dB]	$\Omega_\mathrm{Att}/2\pi$ [Hz]	$\Omega_\mathrm{Amp}/2\pi$ [Hz]
BPF 1	112	13	$[107, 119]$	$[0, 106] \cup [120, f_\mathrm{s}/2]$
BPF 2	141	6	$[138, 142]$	$[0, 139] \cup [143, f_\mathrm{s}/2]$
BPF 3	198	10	$[192, 202]$	$[0, 184] \cup [210, f_\mathrm{s}/2]$
BPF 4	226	6	$[220, 230]$	$[0, 216] \cup [234, f_\mathrm{s}/2]$
BPF 5	375	6	$[368, 383]$	$[0, 367] \cup [384, f_\mathrm{s}/2]$

Right Error Microphone				
	f_p [Hz]	$\eta(\omega_\mathrm{p})$ [dB]	$\Omega_\mathrm{Att}/2\pi$ [Hz]	$\Omega_\mathrm{Amp}/2\pi$ [Hz]
BPF 1	112	13	$[107, 119]$	$[0, 106] \cup [120, f_\mathrm{s}/2]$
BPF 2	198	10	$[192, 202]$	$[0, 184] \cup [210, f_\mathrm{s}/2]$
BPF 3	226	6	$[220, 230]$	$[0, 216] \cup [234, f_\mathrm{s}/2]$
BPF 4	298	10	$[281, 302]$	$[0, 287] \cup [308, f_\mathrm{s}/2]$
BPF 5	375	6	$[368, 383]$	$[0, 367] \cup [384, f_\mathrm{s}/2]$

6.4.3 Simulation Results

The results for the case of 8^{th} ord. and 4^{th} ord. secondary paths are presented in this section.

Simulation Results – 8^{th} ord. Secondary Paths

The secondary paths are employed with the left or right loudspeaker/microphone pairs active one at a time. Magnitude frequency responses of the secondary paths, are depicted in Figs. 6.14 and 6.15. The corresponding simulation parameters are shown in Table 6.2. For all the considered simulations, the quality factor is set to $Q_p = 30$. The length of the filter $\mathcal{K}(z)$ is set to $N = 100$, that, at $f_s = 3$ kHz, corresponds to an impulse response duration of approximately 33 ms.

The SPL power spectra in dBA/Hz for the case of a narrow BPF centered at $f_{p_1} = 112$ Hz, are depicted in Fig. 6.18 for the left (a) and right (b) error microphones. The noise is well mitigated since the green line (ANC ON) is below the red one (ANC OFF), yielding a noise attenuation, around f_{p_1} of 4.28 dB and 4.21 dB for the left

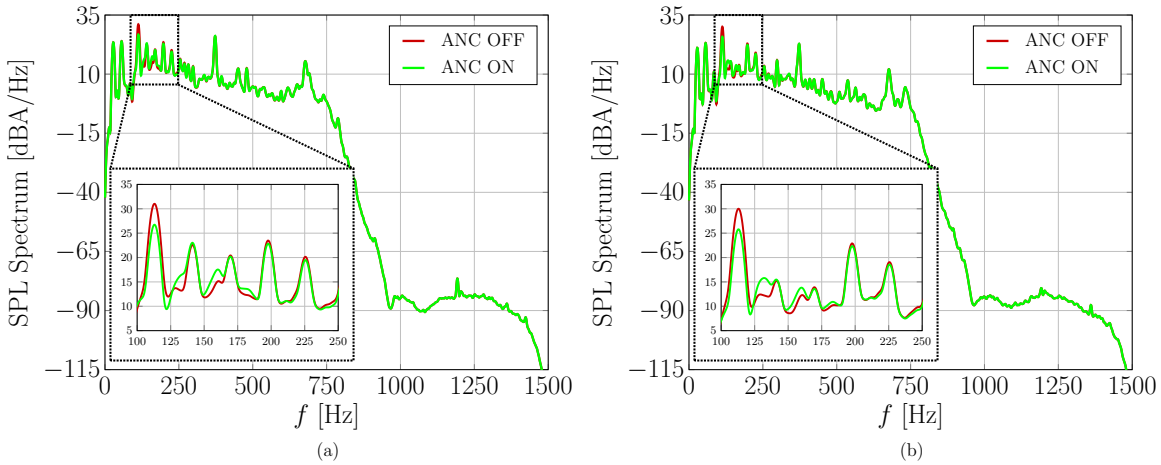


Figure 6.18: SPL power spectra for ANC OFF and fixed SISO FB ANC ON with the hybrid compact car input signal for left (a) and right (b) error microphones. Single-band attenuation at 112 Hz is considered.

and right microphones, respectively. For both microphones, we may observe a slight deterioration around 130 and 160 Hz, which can be considered negligible since its power level is much less than the overall noise level within the cabin. Therefore, it is possible to conclude that the ANC system operates effectively. Same considerations can be drawn by looking at the corresponding sliding window SPL performance depicted in Fig. 6.19. For both error microphones, the noise is nicely mitigated during the simulation time.

In order to implement a multi-band noise cancellation behavior, a set of L_F narrow BPFs can be used. A set of central frequencies are selected based on the peak energy perceived by each microphone, as summarized in Table 6.2.

The SPL power spectra for the multi-band ANC system OFF and fixed SISO ANC ON for the left and right error microphones are depicted in Fig 6.20. The left error microphone exhibits a maximum peak cancellation of 4.25 dB at 197 Hz. Regarding the right error microphone the corresponding maximum noise attenuation is 4.18 dB at 112 Hz. Note that, for the left microphone, cancellation around the peak at 141 Hz does not occur. Moreover, the filter causes a noise amplification around 170 Hz. However for both microphones, performance can be considered satisfactory since the noise peaks are mitigated.

Similar remarks can be made by observing the sliding window SPL in Fig. 6.21 for

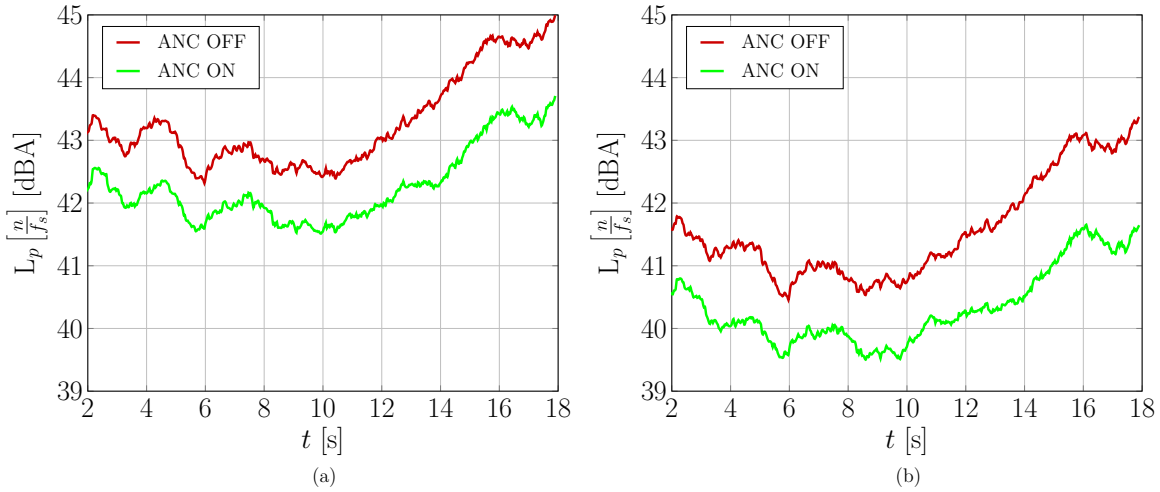


Figure 6.19: Sliding window SPL against time for ANC OFF and fixed SISO FB ANC ON with the hybrid compact car input signal for the left (a) and (b) error microphones. Single-band attenuation at 112 Hz is considered.

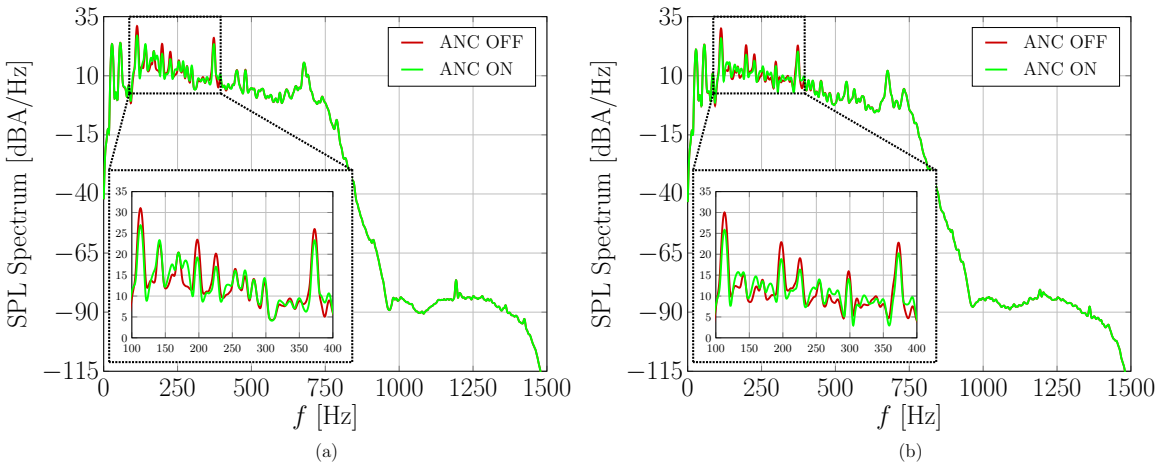


Figure 6.20: SPL power spectra for ANC OFF and fixed SISO FB ANC ON with the hybrid compact car input signal. (a) Left and (b) error microphones. Multi-band attenuation is considered.

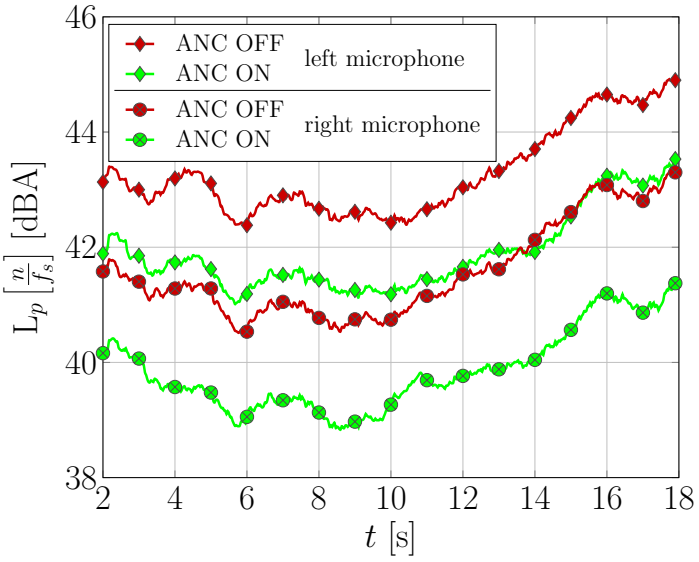


Figure 6.21: Sliding window SPL against time for ANC OFF and fixed SISO FB ANC ON with the hybrid compact car input signal. (a) Left error microphone is diamond-marked and (b) right error microphone is circle-marked. Multi-band attenuation is considered.

both microphones. As previously observed, due to the amplification around 170 Hz, the performance for the left microphone is worse than that for the second one. The right error microphone, in fact, exhibits a maximum and an average cancellation equal to 2.13 dBA and 1.72 dBA, respectively, against 1.85 dBA and 1.38 dBA for the first one.

The SPL spectrum analysis for a SISO adaptive FB ANC system with the hybrid compact car input signal, for the left error microphone, is shown in Fig. 6.22. It is possible to note that at low-frequency, the system is not working since the two curves overlap with each other. This is due to two reasons: first, the magnitude response of the secondary path in that range is very low and second, the A-weighting of the two curves reduces significantly the difference between each other. However, even if we consider the pass-band region, going from 100 to 750 Hz, the system is not working properly since, similarly to what we observed in Fig. 6.5, also here, the cancellation noise, occurring in a certain frequency range, is nullified by an amplification occurring at different frequencies.

Similar considerations can be carried out for the MIMO case shown in Fig. 6.23.

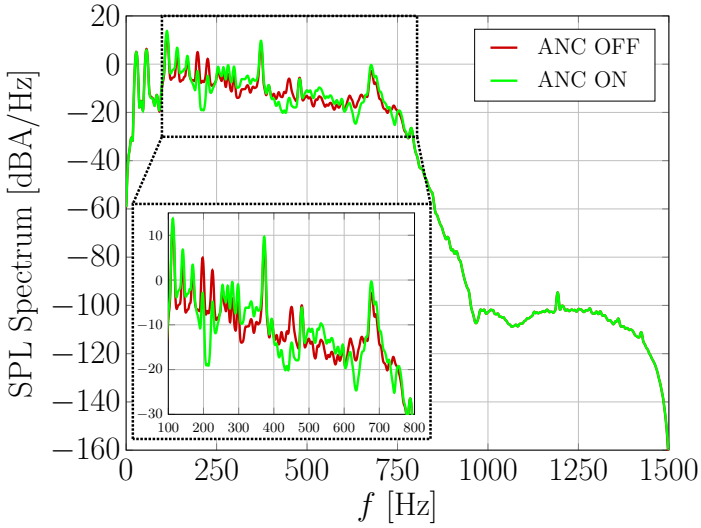


Figure 6.22: SPL spectra with ANC OFF and ANC ON for a SISO adaptive FB ANC system for the left error microphone with the hybrid compact car input signal.

However, by considering the right error microphone, it is possible to observe that three main peaks around about 110 Hz, 195 Hz and 220 Hz are mildly mitigated. The filtered hybrid compact car input signal was employed for the last simulation. In Fig. 6.24, the SPL spectra of the MIMO adaptive FB ANC system for left and right error microphone is shown. As previously mentioned, limiting the bandwidth of the input signal yields some benefits in terms of noise cancellation. In fact, we can note that, for both error microphones the main noise peak components are well mitigated. It is important to highlight that, similarly to what we observed for FF ANC systems in Part I, the FB system is able to cancel the noise peak around 200 Hz by 9 dB and 5 dB, for left and right error microphone, respectively. The other noise components experience noise cancellation from 2 to 4 dB within all the considered bandwidth. Furthermore, noise amplification is negligible since the signal energy is much less than the noise level within the cabin. For such reasons, these results are remarkable and this system may be taken in consideration for a real implementation in the automotive context.

Simulation Results – 4^{th} ord. Secondary Paths

The SPL spectra of MIMO and FE-MIMO FB ANC systems for the left and right error microphones are shown in Fig. 6.25 and compared with the SPL spectra in the

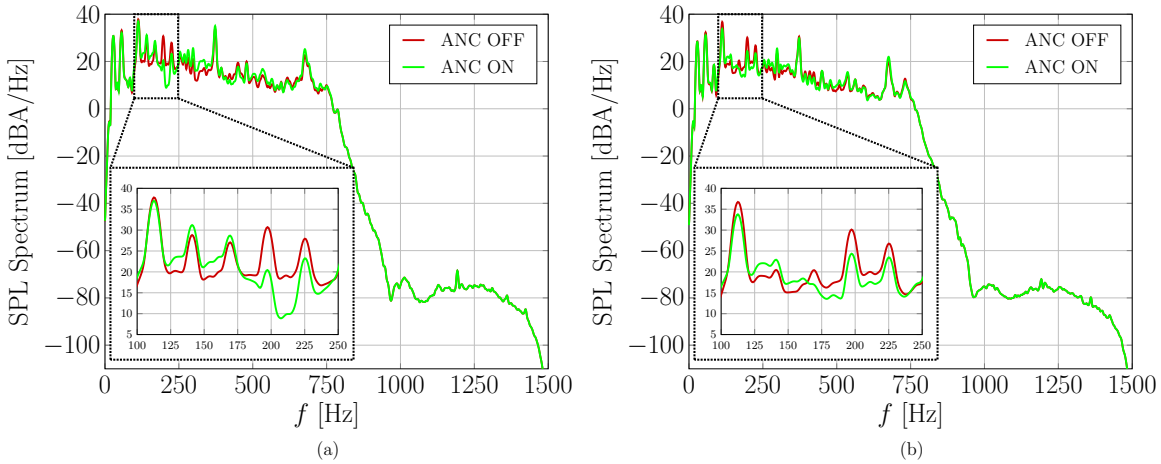


Figure 6.23: SPL spectra with ANC OFF and ANC ON for a MIMO adaptive FB ANC system. (a) Left and (b) right error microphones with the hybrid compact car input signal.

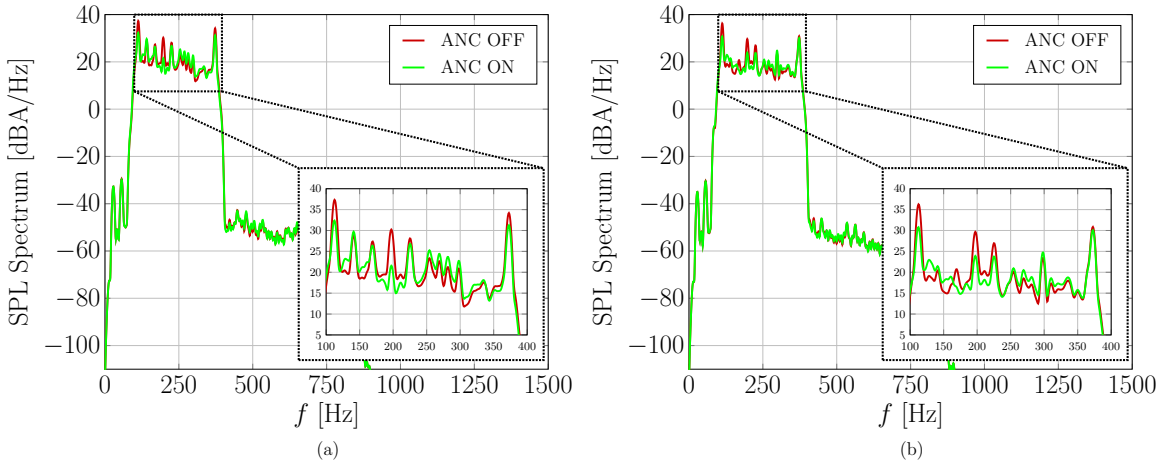


Figure 6.24: SPL spectra with ANC OFF and ANC ON for a MIMO adaptive FB ANC system. (a) Left and (b) right error microphones with the filtered version of the hybrid compact car input signal.

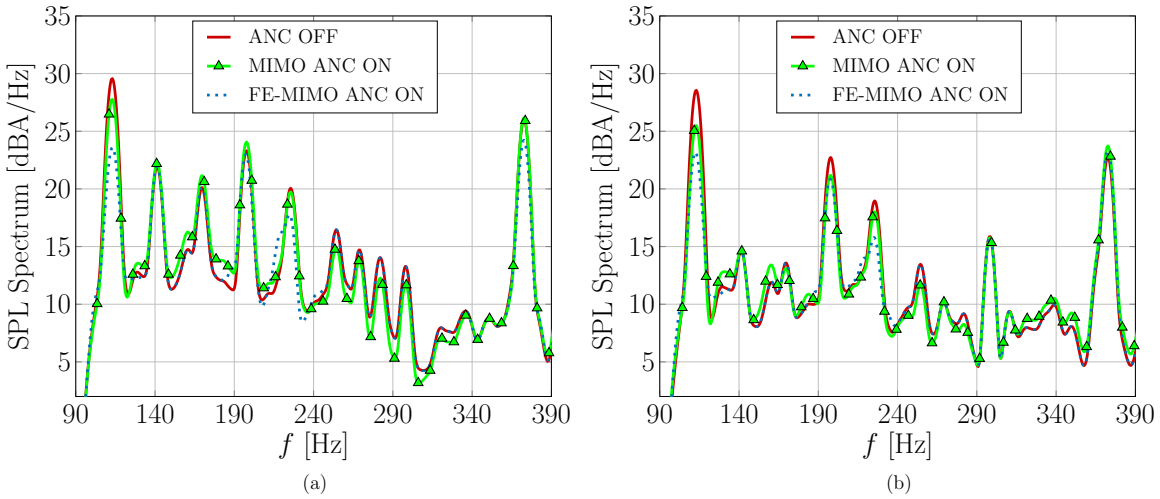


Figure 6.25: SPL power spectra for ANC OFF and ANC ON with standard MIMO and FE-MIMO FB ANC systems. (a) Left and (b) right error microphones.

absence of ANC. Note that the FE-MIMO is the MIMO version of the FB ANC system depicted in Fig. 6.8.

Four frequencies, corresponding to peak power of the disturbance signals, are considered for the filter $V(z)$ of the FE-MIMO system, namely 112 Hz, 197 Hz, 225 Hz and 375 Hz.

It is possible to observe that a band-limited and tonal input signal may facilitate the system operation, showing some benefits in terms of noise cancellation. Especially for the right microphone, the MIMO system (green curve), shows that the main noise peak components, about 112, 200 and 230 Hz, are mitigated up to 2 dB. In these specific frequencies, the performance of the left error microphone is worse with respect to the right one, however, within the range 240–340 Hz, some noise mitigation occurs. Note that in all the considered cases mitigation of the peak around 375 Hz is ineffective due to the the low gain exhibited by the secondary paths, as shown in Figs 6.16 and 6.17. Similar performance is obtained by the FE-MIMO ANC system (dotted blue curve). However, in this case, noise amplification zones are completely avoided and a peak cancellation up to 6 dB is obtained (see the peak at 112 Hz). Only for the left error microphone the peak around 375 Hz is mitigated.

The sliding window SPL performance of the MIMO, FE-MIMO, SISO and ANC system OFF for the left and right error microphones is shown in Fig. 6.26. Note that

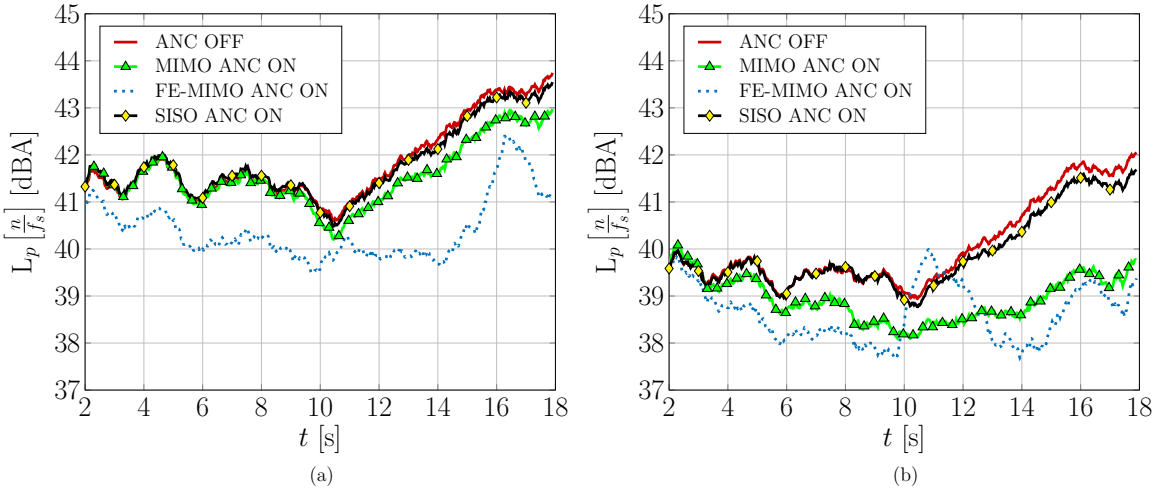


Figure 6.26: Sliding window SPL against time for ANC OFF and ANC ON with standard MIMO, FE-MIMO and SISO FB ANC systems. (a) Left and (b) right error microphones.

the sliding window SPL is calculated as described by (6.1).

Note that the SISO system is obtained with ($M = K = 1$). Therefore, for the left error microphone the employed disturbance signal and the acoustic channel are $d_1[n]$ and $S_{11}(z)$, respectively. Similarly, for the right error microphone they are $d_2[n]$ and $S_{22}(z)$.

The input signal lasts about 18 seconds, but, in order to exclude the initial transient period, the first 2 seconds are discarded. Measurement time therefore lasts about 16 seconds. For both error microphones, the sliding window SPL performance of the MIMO ANC system (triangle-marked green curve) is always better compared to ANC OFF (solid red curve). This behavior shows that, within the considered bandwidth and over the entire simulation time, an appreciable noise cancellation occurs. In particular, the left error microphone exhibits an average and maximum peak noise cancellation of about 0.29 dBA and a 0.79 dBA respectively. System performance is better for the right error microphone which shows an average noise cancellation of about 1.13 dBA and a maximum of 2.42 dBA. Similar performance, with larger gain in terms of noise cancellation, is obtained by the FE-MIMO ANC system (dotted blue curve). More precisely, for the left error microphone, the average and the maximum peak noise cancellation is 1.44 dBA and 2.81 dBA, respectively. Regarding the right error microphone, the

considered values are 1.42 dBA and 3.19 dBA. Finally, for the SISO FB ANC system (diamond-marked black curve), it is possible to conclude that the performance for the left error microphone is not appreciable. Slightly better performance in terms of noise cancellation is obtained for the right error microphone. In particular, the average and the maximum peak noise cancellation is 0.13 dBA and 0.38 dBA, respectively.

Appendix on ANC Systems

A Further Mathematical Derivation of FF System

In this Appendix, the transfer function between the input and the output of the SISO FF ANC system is derived.

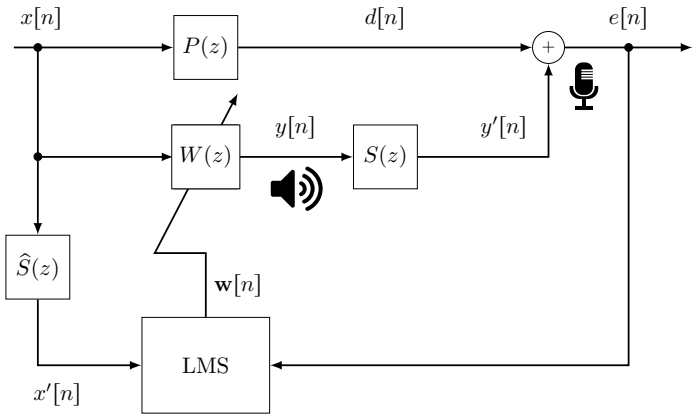


Figure A.1: Block diagram of the single reference SISO FF ANC system.

Let us consider a single-reference FF ANC system as depicted in Fig. A.1. Under ideal condition (i.e., $P(z) = z^0 = 1$), a perfect correlation between the reference signal $x[n]$ and the desired output $d[n]$, i.e., $d[n] = x[n]$, is obtained. The error signal can be written as

$$\begin{aligned} e[n] &= d[n] + y'[n] \\ &= d[n] + x[n] \otimes w[n] \otimes s[n] \\ &= d[n] + d[n] \otimes w[n] \otimes s[n] \\ &= d[n] \otimes (\delta[n] + w[n] \otimes s[n]) \ . \end{aligned} \tag{A.1}$$

By the using $\mathcal{Z}$-transform, (A.1) can be written as

$$E(z) = D(z)\left[1 + W(z)S(z)\right] \ .$$

Hence, the overall transfer function $T(z)$ between the input signal $d[n]$ and the output signal $e[n]$ becomes

$$T(z) = \frac{E(z)}{D(z)} = 1 + W(z)S(z) \ .$$

In order to analyze the behavior of the control filter $W(z)$, from Fig. A.1 the error signal in z-domain is

$$\begin{aligned} E(z) &= D(z) + Y'(z) \\ &= X(z)P(z) + Y(z)S(z) \\ &= X(z)P(z) + X(z)W(z)S(z) \\ &= X(z)\left[P(z) + W(z)S(z)\right] \end{aligned}$$

The idea is to achieve zero level of the error signal. Hence, setting deterministically $E(z) = 0$, one obtains

$$P(z) = -W(z)S(z)$$

and the optimal control filter is

$$W^{\text{opt}}(z) = -\frac{P(z)}{S(z)} \ . \tag{A.2}$$

Equation (A.2) says that, the ideal optimal adaptive filter $W^{\text{opt}}(z)$ has to be simultaneously proportional to the primary path $P(z)$ and inversely proportional to the secondary path $S(z)$.

B Further Numerical Results on FF ANC Systems

For a more exhaustive analysis, further numerical results on the Feed Forward ANC schemes are here reported. This appendix is structured as follows. The SPL spectra and the comparison analysis between theoretical and effective cancellation relative to the SISO, MISO and MIMO FF ANC systems, not considered in the previous sections, are reported in Sections B.1, B.2 and B.3, respectively. Note that only the left error microphone is considered. Similar considerations can be drawn for the right one.

B.1 SISO ANC

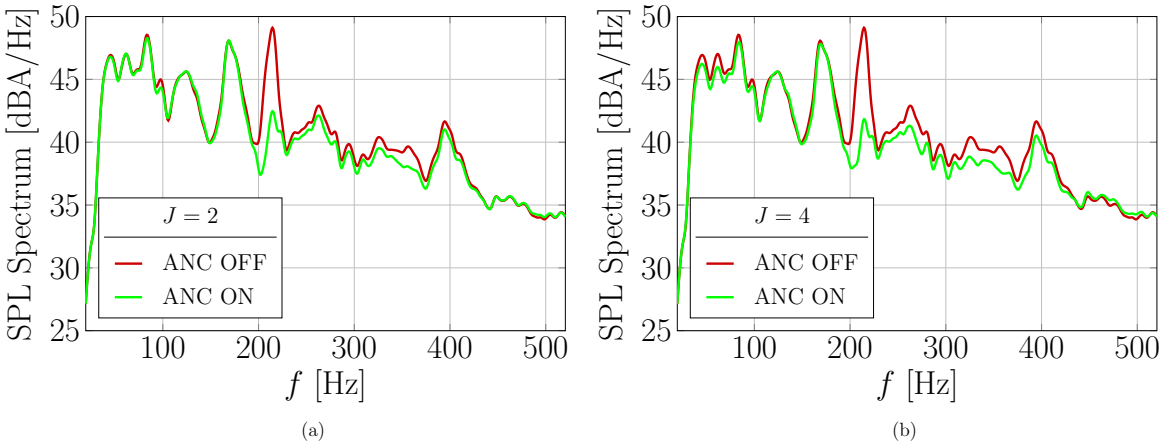


Figure B.1: SPL spectra of $d[n]$ and $e[n]$ in case of SISO ANC system for the best grouping of reference signals with $J = 2$ (a), $J = 4$ (b).

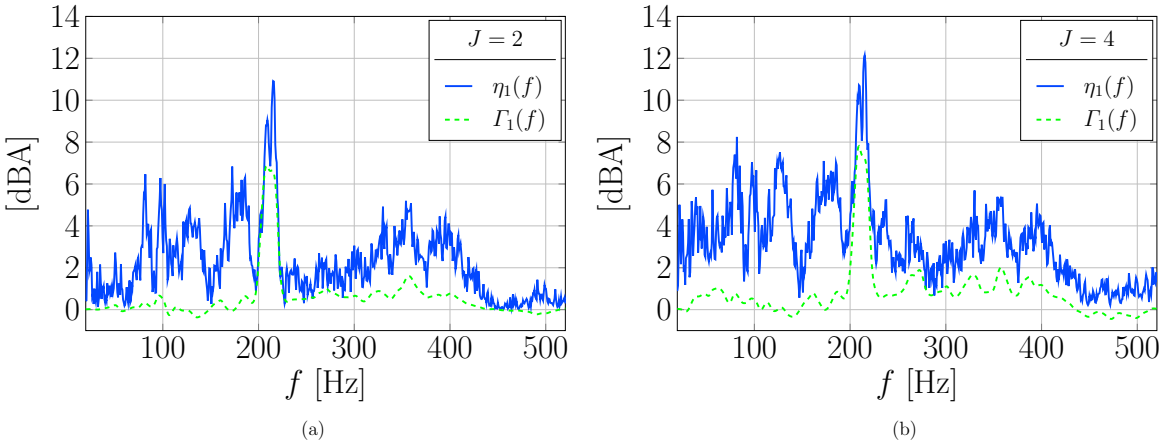


Figure B.2: Comparison between theoretical cancellation $\eta(f)$ and effective cancellation $\Gamma(f)$, as a function of frequency, in case of SISO ANC system for the best grouping of reference signals with $J = 2$ (a), $J = 4$ (b).

B.2 MISO ANC

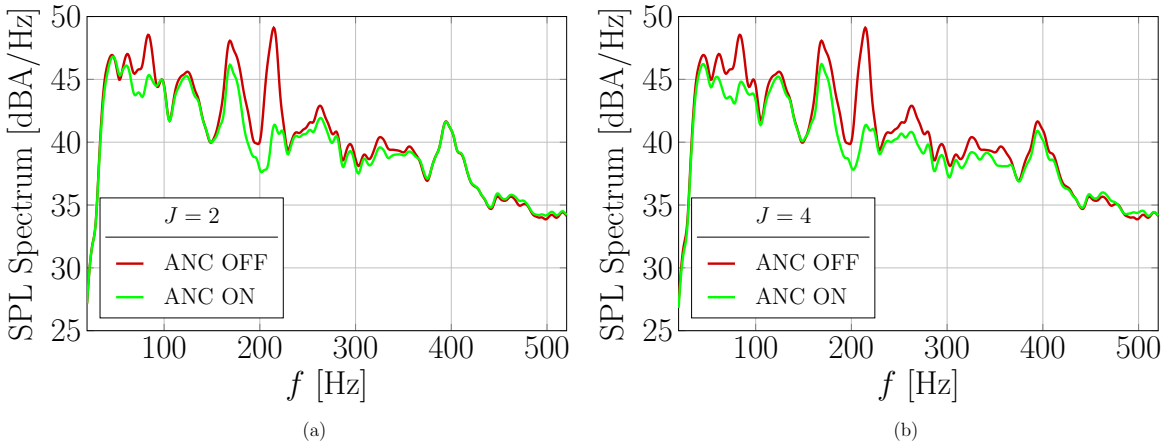


Figure B.3: SPL spectra of $d[n]$ and $e[n]$ in case of MISO ANC system for the best grouping of reference signals with $J = 2$ (a), $J = 4$ (b).

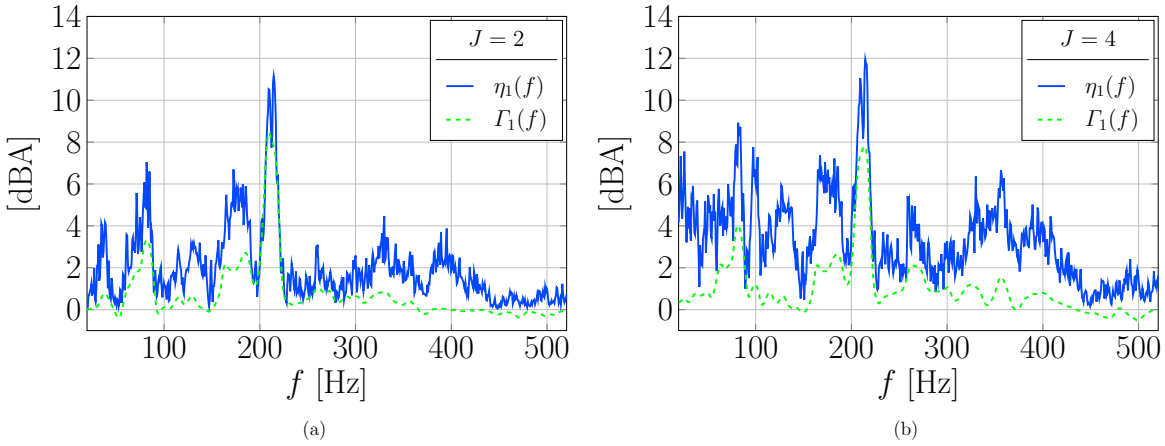


Figure B.4: Comparison between theoretical cancellation $\eta(f)$ and effective cancellation $\Gamma(f)$, as a function of frequency, in case of MISO ANC system for the best grouping of reference signals with $J = 2$ (a), $J = 4$ (b).

B.3 MIMO ANC

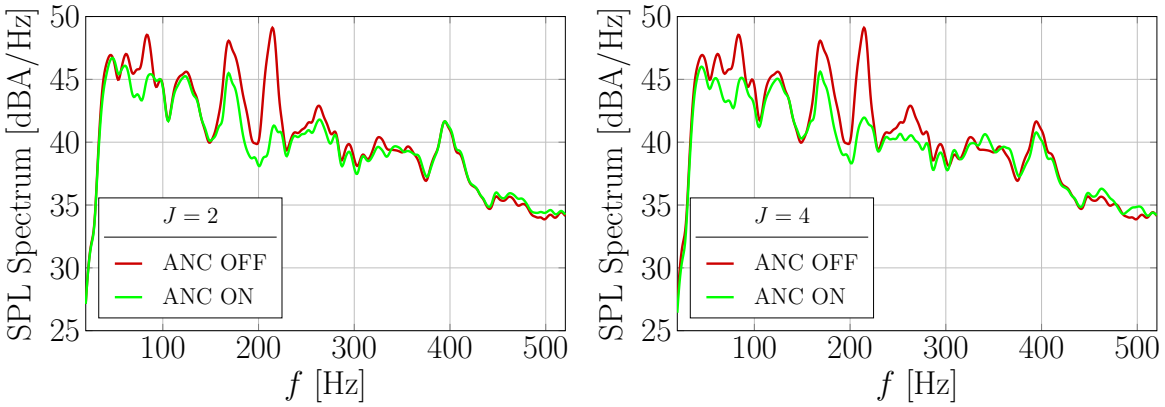


Figure B.5: SPL spectra of $d[n]$ and $e[n]$ in case of MIMO ANC system for the best grouping of reference signals with $J = 2$ (a), $J = 4$ (b).

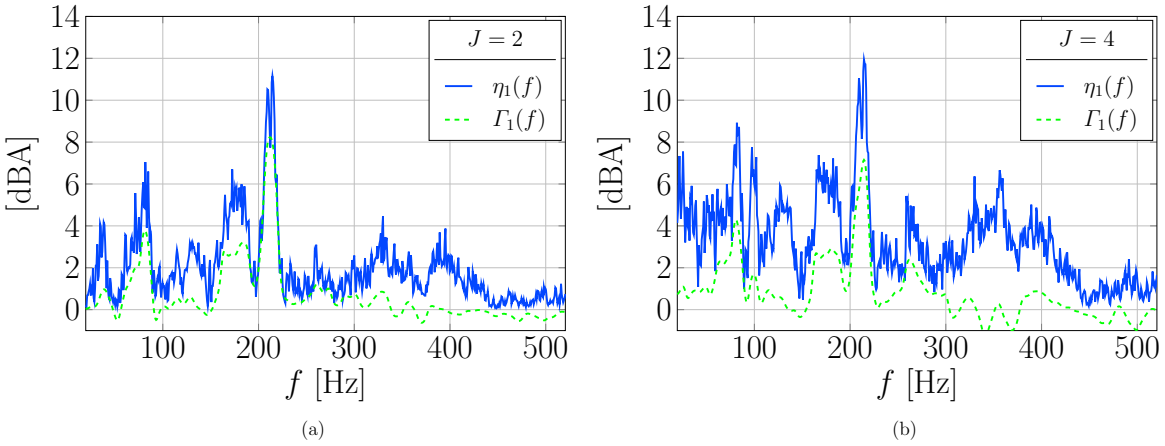


Figure B.6: Comparison between theoretical cancellation $\eta(f)$ and effective cancellation $\Gamma(f)$, as a function of frequency, in case of MIMO ANC system for the best grouping of reference signals with $J = 2$ (a), $J = 4$ (b).

C Further Mathematical Derivation of FB System

The behavior of the system in Fig. 4.2 can be better understood by resorting to an analysis in the $\mathcal{Z}$-domain. For the sake of convenience, the mathematical derivation is given for a SISO system, but can be extended to a MIMO one. The $\mathcal{Z}-$transform of the error microphone signal can be written as

$$E(z) = D(z) + Y'(z) = D(z) + Y(z)S(z) \ . \tag{C.1}$$

From (4.3), one can express the input signal to the loudspeaker as

$$Y(z) = X(z)W(z) \ . \tag{C.2}$$

Thus, by substituting (C.2) into (C.1) and imposing zero error $E(z) = 0$, one gets

$$D(z) + X(z)W(z)S(z) = 0 \ \longrightarrow \ W^{\text{opt}}(z) = -\frac{D(z)}{X(z)S(z)} \tag{C.3}$$

where $W^{\text{opt}}(z)$ denotes the optimal filter. If the secondary path is perfectly estimated, i.e., $\widehat{S}(z) = S(z)$, one obtains

$$X(z) = E(z) - \widehat{Y}'(z) = E(z) - Y(z)\widehat{S}(z) = E(z) - \underbrace{Y(z)S(z)}_{Y'(z)} = D(z) \ . \tag{C.4}$$

Hence, (C.3) becomes

$$W^{\text{opt}}(z) = -\frac{\cancel{D(z)}}{\cancel{X(z)}S(z)} = -\frac{1}{S(z)} \ . \tag{C.5}$$

Eq. (C.5) states that the optimal filter $W^{\text{opt}}(z)$ has to be the inverse of the secondary path. If the secondary path is a pure delay, $W^{\text{opt}}(z)$ is an anti-causal filter anticipating the delay effect of the secondary path, which is clearly not physically realizable.

Under ideal conditions, from (C.4) we have $X(z) = D(z)$. This yields

$$\begin{aligned} E(z) &= D(z) + X(z)W(z)S(z) \\ &= D(z) + D(z)W(z)S(z) \\ &= D(z)\left[1 + W(z)S(z)\right] \ . \end{aligned}$$

The overall transfer function $T(z)$ of the FxLMS ANC system from input signal $d[n]$ to error signal $e[n]$ can be thus expressed as

$$T(z) = \frac{E(z)}{D(z)} = 1 + W(z)S(z) \ .$$

Note that the transfer function of the FB ANC system is equivalent (except for the change of sign) to that of the FF ANC system highlighted in Appendix A.

D Wiener Filter for FB ANC System

In this appendix the Wiener filter solution for fixed FB ANC system is presented. In the first part, the mathematical derivation is presented, then, in the second part the main numerical results are shown and discussed.

Let us consider a SISO fixed FB ANC system as shown in Fig. D.1.

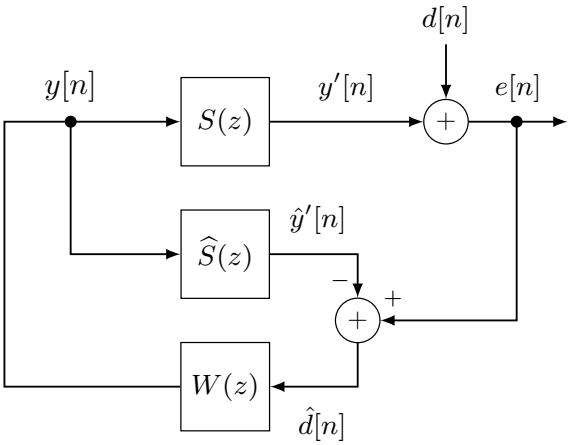


Figure D.1: Block diagram of fixed FB ANC system

The error signal $e[n]$ can be written as

$$\begin{aligned} e[n] &= d[n] + y'[n] \\ &= d[n] + y[n] \otimes s[n] \\ &= d[n] + \hat{d}[n] \otimes w[n] \otimes s[n] \end{aligned} \tag{D.1}$$

where the reconstructed reference signal $\hat{d}[n]$ is equal to

$$\begin{aligned} \hat{d}[n] &= e[n] - \hat{y}'[n] \\ &= e[n] - y[n] \otimes \hat{s}[n] \\ &= e[n] - \hat{d}[n] \otimes w[n] \otimes \hat{s}[n] \ . \end{aligned} \tag{D.2}$$

By passing to the $\mathcal{Z}$-domain, (D.1) and (D.2) become

$$E(z) = D(z) + \widehat{D}(z)W(z)S(z) \tag{D.3}$$

and

$$\widehat{D}(z) = E(z) - \widehat{D}(z)W(z)\widehat{S}(z) \tag{D.4}$$

respectively. Solving, (D.4) becomes

$$\widehat{D}(z) = \frac{E(z)}{1 + W(z)\widehat{S}(z)} \ . \tag{D.5}$$

Now, by plugging (D.5) into (D.3) one gets

$$E(z) = D(z) + \frac{E(z)}{1 + W(z)\widehat{S}(z)} W(z)S(z) \ . \tag{D.6}$$

Finally, rearranging (D.6), the input/output transfer function $G(z)$ can be expressed as

$$T(z) = \frac{E(z)}{D(z)} = \frac{1 + W(z)\widehat{S}(z)}{1 + W(z)\left(\widehat{S}(z) - S(z)\right)} \ . \tag{D.7}$$

Under the assumption of perfect estimation of the secondary path, i.e, $\widehat{S}(z) = S(z)$, (D.7) reduces to

$$T(z) = 1 + W(z)S(z)$$

hence the error signal can be written as

$$E(z) = D(z) + D(z)W(z)S(z) \ . \tag{D.8}$$

Through the inverse $\mathcal{Z}$-transform, (D.8) can be expressed as

$$\begin{aligned} e[n] &= d[n] + d[n] \otimes w[n] \otimes s[n] \\ &= d[n] + \left(\mathbf{w}^\top \mathbf{d}[n]\right) \otimes s[n] \\ &= d[n] + \left(\mathbf{d}^\top[n]\mathbf{w}\right) \otimes s[n] \end{aligned} \tag{D.9}$$

where the length-N column vectors of the desired signal and tap weights are defined as $\mathbf{d}[n] = [d[n], d[n-1], ..., d[n-N+1]]^\top$ and $\mathbf{w} = [w_0, w_1, ..., w_{N-1}]^\top$, respectively.

Assuming a pure delay of n_0 samples as secondary path, the impulse response $s[n]$ can be written as

$$s[n] = \delta[n - n_0]$$

being $\delta[\cdot]$ the Kronecker-delta. Hence, (D.9) becomes

$$\begin{aligned} e[n] &= d[n] + \mathbf{w}^\top \mathbf{d}[n - n_0] \\ &= d[n] + \mathbf{d}^\top[n - n_0]\mathbf{w} \ . \end{aligned}$$

Similarly to Subsection 1.1, a cost function χ can be defined and evaluated as follows

$$\begin{aligned} \chi &= \mathbb{E}\left\{e^2[n]\right\} \\ &= \mathbb{E}\left\{e[n]e[n]\right\} \\ &= \mathbb{E}\left\{\left[d[n] + \mathbf{d}^\top[n - n_0]\mathbf{w}\right]^\top \left[d[n] + \mathbf{d}^\top[n - n_0]\mathbf{w}\right]\right\} \\ &= \underbrace{\mathbb{E}\left\{d^2[n]\right\}}_{\sigma_d^2} - 2\underbrace{\mathbb{E}\left\{d[n]\mathbf{d}^\top[n - n_0]\right\}}_{\mathbf{b}^\top}\mathbf{w} + \mathbf{w}^\top \underbrace{\mathbb{E}\left\{\mathbf{d}[n - n_0]\mathbf{d}^\top[n - n_0]\right\}}_{\mathbf{A}}\mathbf{w} \\ &= \sigma_d^2 - 2\mathbf{b}^\top\mathbf{w} + \mathbf{w}^\top\mathbf{A}\mathbf{w} \ . \end{aligned}$$

Now, by setting the gradient of the cost function $\nabla_{\mathbf{w}}(\chi)$ equal to zero, the *Wiener-Hopf* solution for the considered FB scheme is equal to

$$\boxed{\mathbf{w}_0 = -\mathbf{A}^{-1}\mathbf{b} \ .}$$

In particular, we can express the length-N vector $\mathbf{b}$ as

$$\mathbf{b} = \mathbb{E}\left\{d[n]\mathbf{d}[n - n_0]\right\} = \begin{bmatrix} r[n_0] \\ r[n_0 - 1] \\ \vdots \\ r[n_0 - N + 1] \end{bmatrix}$$

where $r[m] = \mathbb{E}\{d[n]d[n-m]\}$ and the $N \times N$ matrix $\mathbf{A}$ as

$$\mathbf{A} = \mathbb{E}\left\{\mathbf{d}[n-n_0]\mathbf{d}^\top[n-n_0]\right\} = \begin{bmatrix} r[0] & r[1] & \cdots & r[N-1] \\ r[-1] & r[0] & \cdots & r[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ r[-N+1] & r[-N+2] & \cdots & r[0] \end{bmatrix}.$$

The elements of vector $\mathbf{b}$ and matrix $\mathbf{A}$ can be estimated as a temporal correlation on a window from the index N_0 to N_1. Therefore, $r[m]$ can be expressed as

$$r[m] = \mathbb{E}\{d[n]d[n-m]\} \simeq \frac{1}{N_1 - N_0} \sum_{n=N_0}^{N_1-1} d[n]d[n-m]$$

for $0 \leq m \leq N-1$, $N < N_0$ and $N_1 \leq N_B$, where N_B is the last sample in the available sequence. Note that, by symmetry, one has $r[-m] = r[m]$.

Numerical Results

In this section, the numerical results of the fixed FB Wiener solution are shown.

The SPL spectra for a SISO Wiener FB ANC system with the sedan input signal, for the error microphone 1, are shown in Fig. D.2. Note that the green curve, representing the SPL spectrum for the ANC system ON, is almost flat over the considered frequency range.

A comparison in terms of Wiener filter length N is shown in Fig. D.3(a). It is possible to observe that the system performance is almost the same regardless the filter length.

Finally, a comparison between the fixed approach performed by the Wiener filter and an adaptive approach performed by LMS algorithm, is shown in Fig. D.3(b). We can note that even for several step-size values, the Wiener solution is always optimal with respect to the LMS one if we restrict to the frequency range of interest.

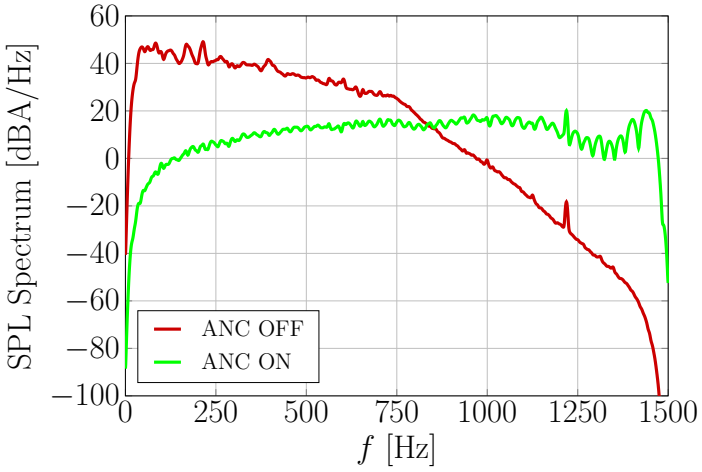


Figure D.2: SPL spectra with ANC OFF and ANC ON for a SISO Wiener FB ANC system for error microphone 1 with the sedan input signal.

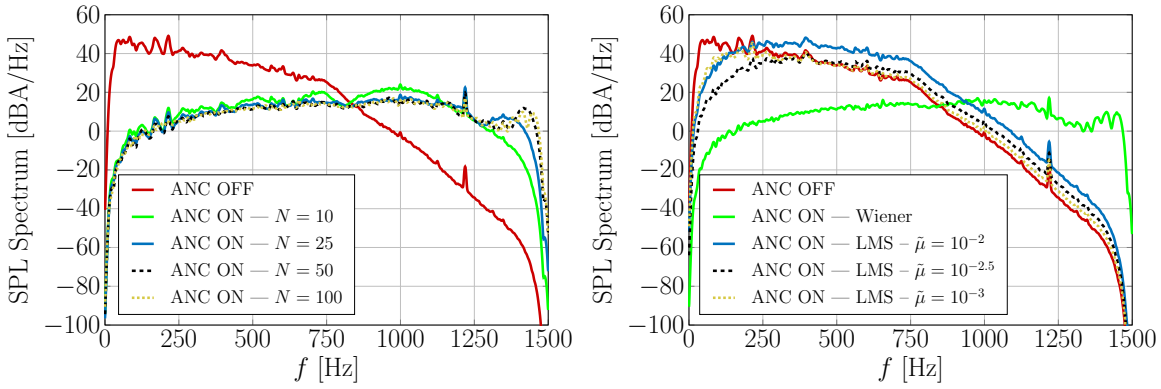


Figure D.3: SPL spectra with ANC OFF and ANC ON for different filter length for a SISO Wiener FB ANC system for error microphone 1 with the sedan input signal (a). SPL spectrum comparison between fixed ANC system (Wiener filter) and adaptive ANC system (LMS algorithm) for error microphone 1 in case of sedan input signal (b).

E Further Numerical Results on FB ANC Systems

In this appendix, further numerical results on adaptive FB ANC systems are presented. In particular, simulation results for the SISO and MIMO ANC systems are shown in Section E.1 and E.2, respectively. Finally in Section E.3, the so-called Combined Input-Combined Output (CICO) adaptive FB ANC scheme is derived and system performance is compared with respect to the MIMO one.

Note that the sliding window SPL (in dBA) for a signal $p[n]$ was defined in (6.1) and is reported here for convenience

$$\mathrm{L}_p[n] = 10 \log_{10} \frac{1}{L p_0^2} \sum_{\ell=1}^{L} p^2[n-\ell]$$

where $p_0 = 20$ μPa, and A-weighting filtering is considered. In our simulations, $L = 3000$ samples, which corresponds to 1 second at $fs = 3$ kHz.

E.1 Adaptive SISO ANC

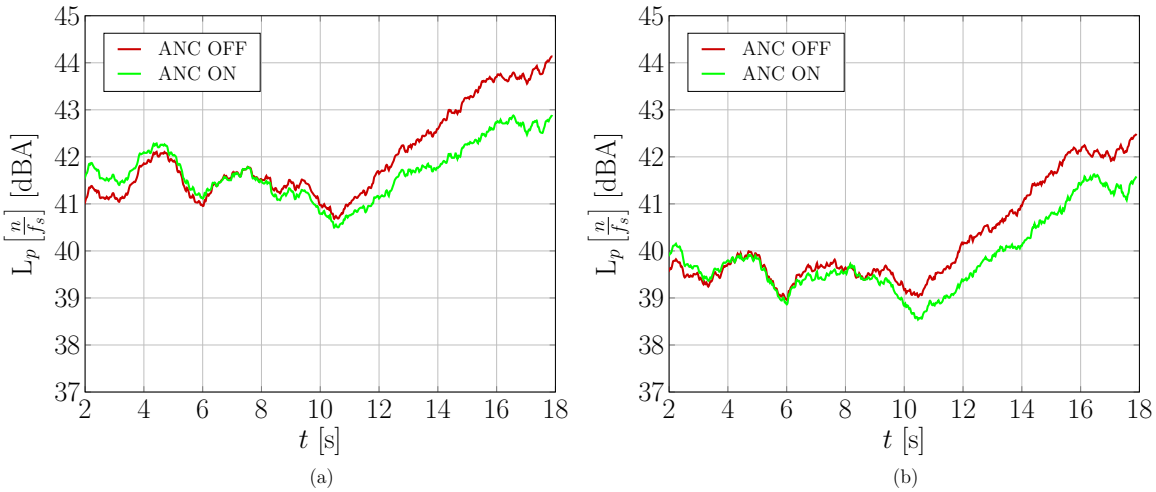


Figure E.4: Sliding window SPL against time for a SISO adaptive FB ANC system. (a) Left and (b) right error microphones with a filtered version of hybrid compact car input signal and sedan as secondary path.

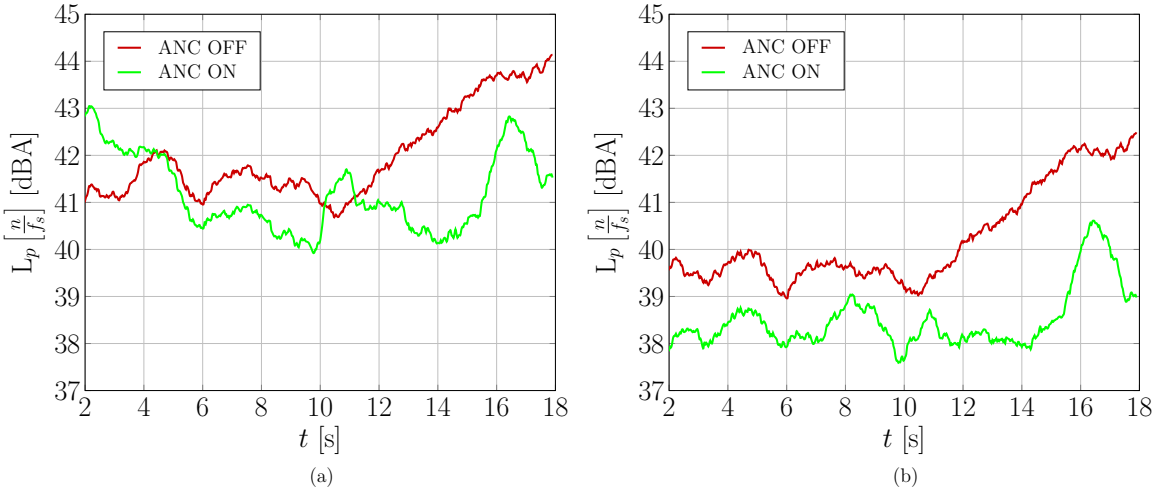


Figure E.5: Sliding window SPL against time for a SISO adaptive FE FB ANC system. (a) Left and (b) right error microphones in case of one filter, a filtered version of hybrid compact car input signal and sedan secondary path.

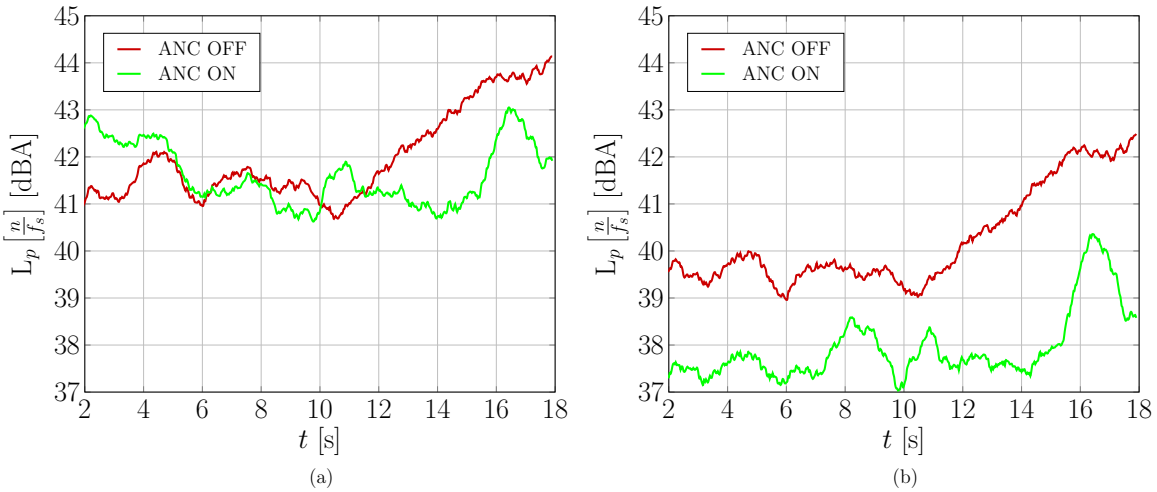


Figure E.6: Sliding window SPL against time for a SISO adaptive FE FB ANC system. (a) Left and (b) right error microphone in case of two filters, a filtered version of hybrid compact car input signal and sedan secondary path.

E.2 Adaptive MIMO ANC

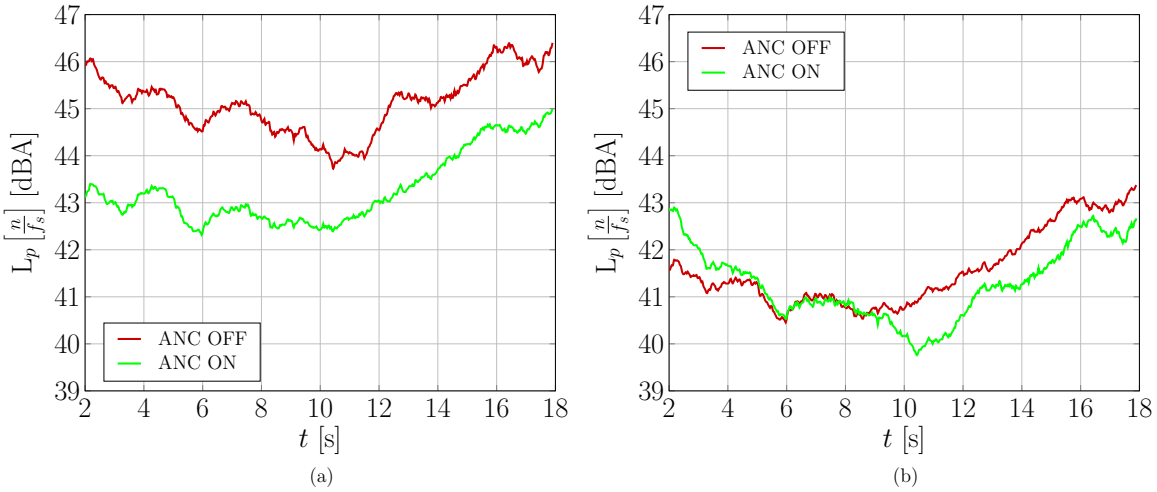


Figure E.7: Sliding window SPL against time for a MIMO adaptive FB ANC system. (a) Left and (b) right error microphones with the hybrid compact car input signal and sedan as secondary path.

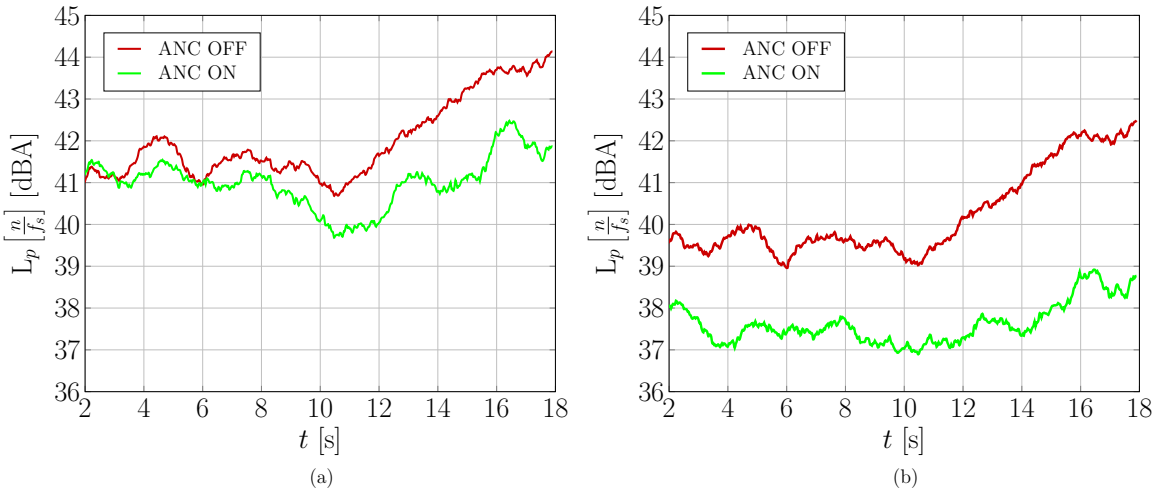


Figure E.8: Sliding window SPL against time for a MIMO adaptive FB ANC system. (a) Left and (b) right error microphones with a filtered version of hybrid compact car input signal and sedan as secondary path.

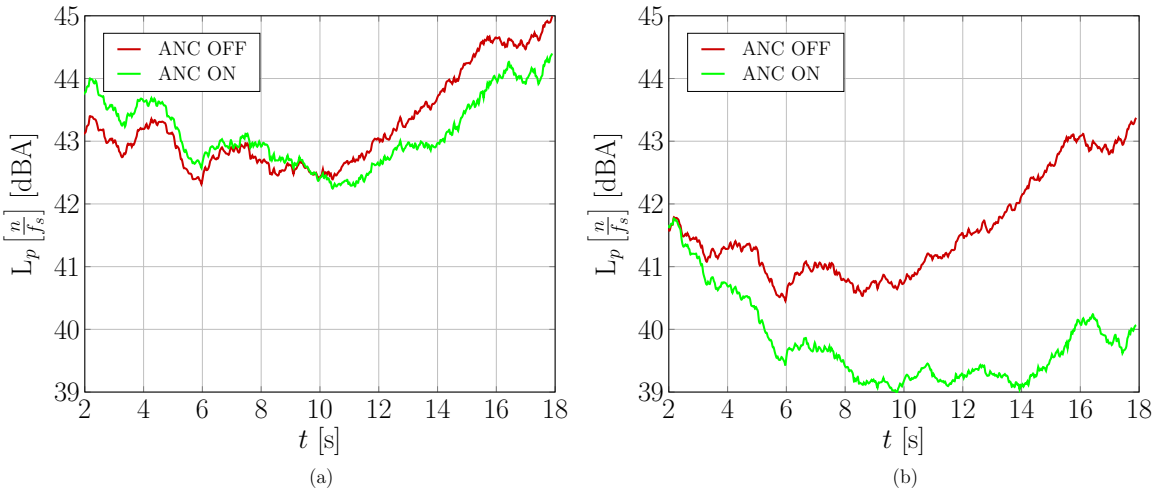


Figure E.9: Sliding window SPL against time for a MIMO adaptive FB ANC system. (a) Left and (b) right error microphones with the hybrid compact car input signal and headrest chair as secondary path.

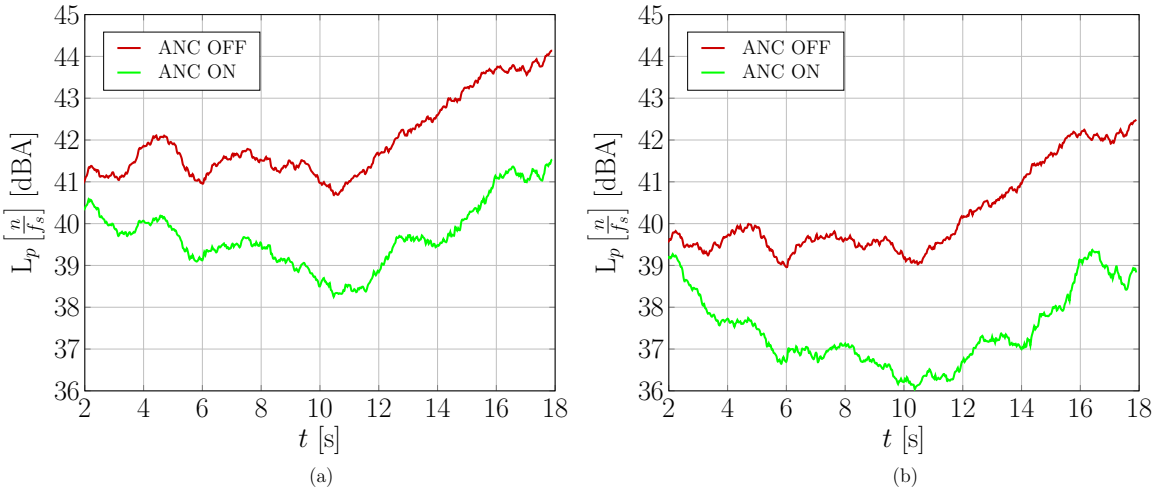


Figure E.10: Sliding window SPL against time for a MIMO adaptive FB ANC system. (a) Left and (b) right error microphones with a filtered version of hybrid compact car input signal and headrest chair as secondary path.

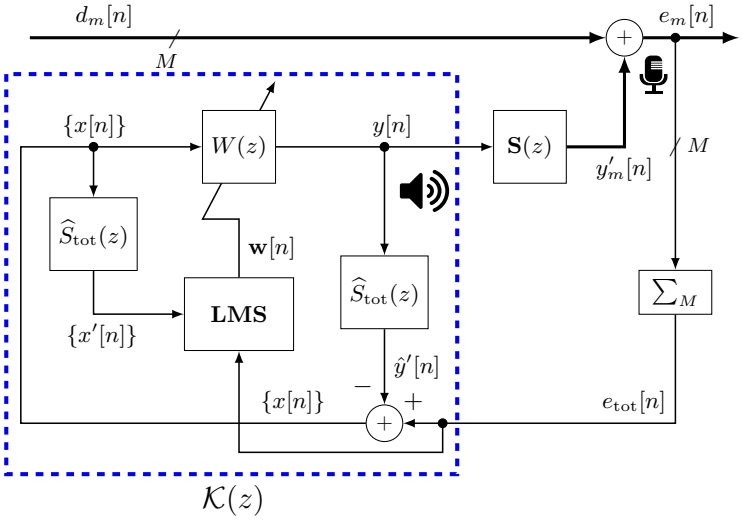


Figure E.11: Block diagram of the adaptive CICO FB ANC system.

E.3 Adaptive CICO ANC

In this section, a special case of the FB MIMO scheme, the CICO system is discussed. This system, considered in [38], was presented as an effective alternative to MIMO FB system, since it guarantees a significant saving in terms of the system complexity, while maintaining good noise cancellation performance. In particular, in [38] an ANC system for office environment is considered. For this purpose, the headrest of an office chair, equipped with two loudspeakers and two error microphones, is employed in order to obtain a quiet zone around the head of the seated person.

In Fig. E.11, a block diagram of the adaptive CICO FB ANC system is depicted. The main difference with respect to the scenarios previously considered is the error signal which feeds the LMS algorithm. In a CICO system, in fact, instead of having K different signals, only one signal feeds the K loudspeakers. The LMS algorithm is therefore driven by $e_{\text{tot}}[n]$, which is equal to

$$\begin{aligned} e_{\text{tot}}[n] &= \sum_{m=1}^{M} e_m[n] \\ &= \sum_{m=1}^{M} \left(d_m[n] + y'_m[n] \right) \ . \end{aligned} \tag{E.10}$$

For each error microphone, the error signal can be expressed as

$$\begin{aligned} e_m[n] &= d_m[n] + y'_m[n] \\ &= d_m[n] + \sum_{k=1}^{K}\sum_{h=0}^{H-1} s_{mk}[n-h] \sum_{\ell=0}^{N-1} x[n-\ell] w_\ell[h] \ . \end{aligned}$$

From (E.10), the synthesized reference signal $x[n]$ is

$$\begin{aligned} x[n] &= e_{\text{tot}}[n] - \sum_{m=1}^{M} \hat{y}'_m[n] \\ &= e_{\text{tot}}[n] - \underbrace{\sum_{m=1}^{M}\sum_{k=1}^{K} \hat{s}_{mk}[n]}_{s_{\text{tot}}[n]} \otimes y[n] \\ &= e_{\text{tot}}[n] - y[n] \otimes \hat{s}_{\text{tot}}[n] \end{aligned}$$

where $\hat{s}_{\text{tot}}[n]$ identifies the estimate of the overall impulse response from the loudspeakers to the combination of error microphones. Similarly to the MIMO case, the input signal of the LMS algorithm is the filtered version of the reference signal $x[n]$ such that

$$x'[n] = x[n] \otimes \hat{s}_{\text{tot}}[n] \ .$$

This means that, for a CICO FB ANC system, the leaky normalized FxLMS update equation becomes

$$\boxed{\mathbf{w}[n+1] = \lambda \mathbf{w}[n] - \tilde{\mu} e_{\text{tot}}[n] \frac{\mathbf{x}'[n]}{\epsilon + \mathbf{x}'^{\top}[n]\mathbf{x}'[n]}}$$

being

$$\mathbf{x}'[n] = \begin{bmatrix} x'[n] \\ x'[n-1] \\ \vdots \\ x'[n-N+1] \end{bmatrix} = \begin{bmatrix} \hat{s}_{\text{tot}}[n] \otimes x[n] \\ \hat{s}_{\text{tot}}[n] \otimes x[n-1] \\ \vdots \\ \hat{s}_{\text{tot}}[n] \otimes x[n-N+1] \end{bmatrix} \ .$$

For the sake of clarity, Fig. E.12 shows a diagram in which the block representing

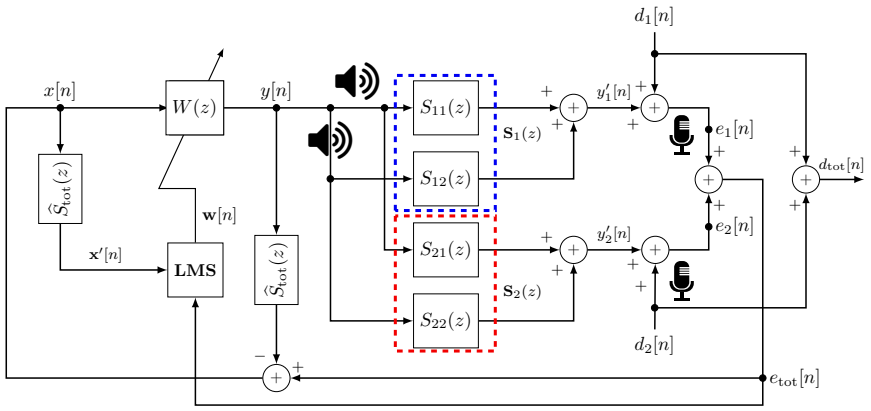


Figure E.12: Expanded block diagram of the adaptive CICO FB ANC system in the case with $M = 2$.

the secondary path $S_{\text{tot}}(z)$ is expanded for a 2×2 CICO FB ANC system. It is possible to observe that, similarly to the MIMO case, a CICO system still employs M error microphones and K loudspeakers, but the loudspeakers are fed by the same signal $y[n]$ and the LMS is driven by e_{tot} instead of $e_m[n]$. This means that the CICO system is equivalent to a SISO system trying to cancel the signal $d_1[n] + d_2[n]$ be means of the overall error signal $e_1[n] + e_2[n]$. The equivalent SISO representation of a CICO system is depicted in Fig. E.13.

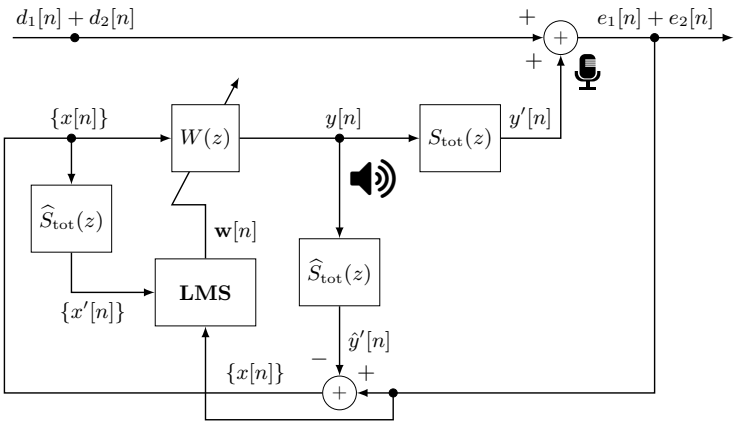


Figure E.13: Equivalent SISO representation of a CICO FB ANC system.

Adaptive MIMO ANC vs adaptive CICO ANC

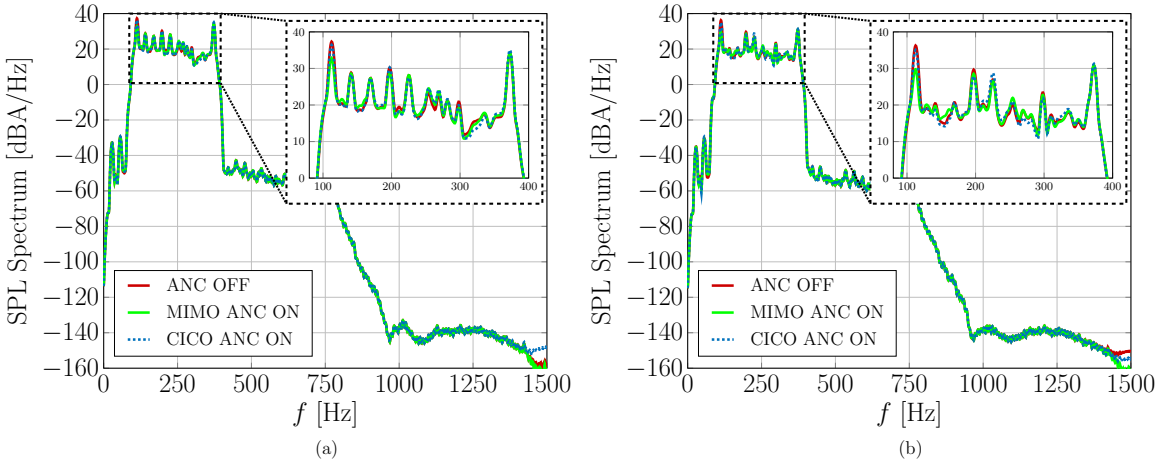


Figure E.14: SPL spectrum comparison between adaptive MIMO FB ANC system and adaptive CICO FB ANC system. (a) Left and (b) right error microphones with the filtered version of hybrid compact car input signal and sedan as secondary path.

Conclusions on ANC Systems

In Part I, a thorough investigation on FF and FB ANC schemes in automotive applications is presented.

The goal of the ANC system is to reduce the perceived audio noise by properly estimating the primary path between the reference signals and the error microphones. Starting with a mathematical derivation of the FF Wiener filter in Chapter 1, the multiple reference MIMO FxLMS-based solution is formulated in Chapter 2. The simulations are performed using MATLAB and Simulink with realistic measurements. Three types of multi-reference FF ANC systems are designed and analyzed, depending on the considered numbers of loudspeakers and error microphones, namely SISO (Section 3.2.1), MISO (Section 3.2.2), and MIMO (Section 3.2.3). Six reference signals are employed with the best groupings, which show the best average performance, are discussed. Obtained results are assessed in terms of SPL, both for the desired output and error signals, effective cancellation (also compared with the theoretical one), and spectrograms when ANC system is ON and OFF. Results show how the main noise component around 200–220 Hz, due to the cavity modes of a rolling tire, is properly canceled in all the considered scenarios. Moreover, the MISO scenario allows us to cancel noise in a wider spectrum with respect to the SISO one due to the inherent system diversity. Finally, the MIMO system is able to control the perceived noise at more than one position, at the expense of reduced peak noise cancellation.

Chapter 4 is dedicated to the mathematical derivation for adaptive and fixed FB ANC system. In particular, both for adaptive and fixed FB ANC system, three types of schemes are implemented, i.e., Standard FxLMS and Modified FxLMS and in Section 4.1 and LPF, NBPF and BBPF in Section 4.2, respectively. Chapter 5 is dedicated to the presentation of the considered input signals for the FB ANC systems. Numerical results, shown in Chapter 6, prove that the main limit of FB ANC system is the delay

introduced by the secondary paths. To cope with this issue an experimental headrest chair is set-up as presented in Section 6.4. The reduced distance between loudspeakers and microphones improves the system performance. Furthermore, it is shown that the FB ANC system exhibits performance improvements for a very narrow-band or tonal signal disturbance. However, in case of adaptive FB MIMO ANC system with filtered hybrid compact car input signal, results are quite encouraging since the main noise components are canceled.

Part II

Virtual Microphone Technique

Introduction

Part I of this thesis is dedicated to a deep investigation on the performance of FF and FB ANC systems for the case of SISO, MISO and MIMO schemes. In particular, the anti-noise signals, emitted by the loudspeakers, are synthesized by employing both the adaptive approach, i.e., the LMS algorithm and the fixed one.

Although these systems, for several reasons, are different from each other, they share the same operating goal: cancel the disturbance noise at the error microphone level. If on one hand, the error microphone positioning can be optimized in order to maximize the noise cancellation performance, in the automotive industry, microphone placing is usually constrained by the car producer, so that physically reaching the desired quiet zone may be very complicated. It is therefore possible to distinguish between the physical error microphone that is employed to monitor to the incoming noise, usually referred to as *monitoring microphone*, and the *virtual microphone*, that is the physical region in which the noise cancellation is targeted, e.g., in our case, near the driver/passenger's ears.

In the literature, this approach is known as Virtual Microphone Technique (VMT) also known as Remote Microphone Technique (RMT) [39–41]. The idea behind of VMT is to reconstruct the virtual microphone signals by exploiting the monitoring ones. In principle, in fact, when ANC with VMT is employed, it is possible to minimize the disturbing audio waves at the virtual locations. To do this, the accurate estimation of the physical acoustic channels between monitoring and virtual microphones, usually referred to as Observation Filter (OF), is needed. It is expected that, the larger the number of monitoring microphones employed to virtualize a position, the better the observation filter estimation accuracy. In the literature several algorithms are proposed [39, 42] and employed in the automotive context [43, 44]. The performance investigation on the accuracy of the OF estimation, when single and several monitoring

microphones are employed to virtualize single and multiple positions in realistic and experimental automotive scenarios, represents the topic of this Part II.

The rest of this part is organized as follows. The mathematical background on the ANC algorithms with OF are presented in Chapter 7. Chapter 8 is dedicated to the adopted algorithms for the OF estimation.

The synthetic and experimental measurements for virtual microphone acquisition in realistic conditions and their corresponding experimental results are shown in Chapter 9. In Chapter 10, a new possible approach to improve the current state of the art in microphone virtualization, based on Temporal Convolutional Network (TCN), is studied. Finally, concluding and remarks are drawn.

State of Art

Microphone virtualization, also known in the literature as VMT, part of a general family of Virtual Sensing Techniques [45, 46], is used to tackle the problem of attenuating acoustic noise in scenarios in which the monitoring microphone cannot be placed at the target quiet zone location. For example, for ANC applications in automotive scenarios, the users are preferably not required to wear headsets, therefore the monitoring microphones have to be placed in locations near the target quiet zone. Additionally, employing beam-forming algorithms [47], it is possible to configure the system to make use of multiple monitoring microphones placed around multiple target quiet zones [48].

Microphone virtualization is usually based on the extraction of observation filters, used to process the audio captured by physical microphones to obtain the estimated audio propagating in the target quiet zone, which is then fed to the ANC system for noise cancellation. However, observation filters have to be extracted off-line, and therefore they are only able to model stationary channels. For this reason, the system may be sub-optimal if the actual channel is time-varying. In particular, in the automotive scenario, the channel may change depending on the user head movement, and on some environmental driving conditions, e.g., the presence of a passenger.

Additionally, another limitation of microphone virtualization is the spatial aliasing that affects the system at high frequencies as the distance increases. The virtual microphone accuracy is bound by a distance-bandwidth trade off, since the higher frequencies of the sound at the target quiet zone are hard to predict if the monitoring microphones are too far. In order to obtain the highest accuracy, the monitoring microphones have to be placed the closest possible to the target quiet zone. In a recent work, these issues were successfully avoided by substituting the monitoring microphones with laser doppler vibrometers (LDVs), which precisely capture the vibrations at the user ear [49]. This optical microphone setup is not bound by distance since the laser beam captures

the vibration directly on the surface of the ear. However, this setup is rather expensive and requires the use of reflective tape to be placed on the ear, which is comparable to the case of requiring the user to wear a headset or headphones.

Solutions based on the geometric arrangement of physical and virtual microphones, can be found in [50]. Similarly to the estimation of the observation filter, that requires a training stage in which physical transducers have to be placed at the desired quiet zone, the use of the so-called *additional filter*, was proposed in [51,52]. The idea behind this method is to perform the ANC at the desired quiet zone, i.e., minimize the virtual error microphone signal, and to train an additional filter in order to identify (to train) the acoustic paths between the reference signals and the monitoring microphones once the virtual error microphone signals are minimized. Then, during the control phase, the estimated additional filter is kept fixed and the FxLMS minimizes the difference signal between the monitoring microphone and the additional filter output.

Chapter 7

ANC with VMT

The goal of this chapter is to describe the operating scheme of the ANC system with VMT. In particular, the analysis is focused on the fixed and adaptive control filter derivation for the anti-noise signal in Sections 7.1 and 7.2, respectively. The observation filter is assumed to be known.

Before the mathematical derivations, in this context, it is necessary to make some clarifications on the adopted nomenclature. In general, in order to virtualize one virtual error microphone, one or more than one monitoring (physical) error microphones can be employed. Consequently, it is also possible to use one single monitoring microphone for several virtual microphones or several monitoring microphones in order to virtualize more than one virtual error microphones. System performance, in terms of estimation accuracy, increases when several physical microphones are employed in order to virtualize a fixed number of virtual positions. Hence, in the context of microphone virtualization, the acronyms SISO, MISO and MIMO will be used to identify the number of loudspeakers (input) and the number of physical monitoring microphones (output). The number of virtual microphones, instead, will be specified in terms of number of quiet zones, i.e., one quiet zone for one virtual error microphone, multiple quiet zones for several virtual error microphones.

For the sake of simplicity, the following mathematical background is only presented for the case of a single quiet zone, single reference, SISO ANC system. However, the derivations can be easily extended to MISO and MIMO systems.

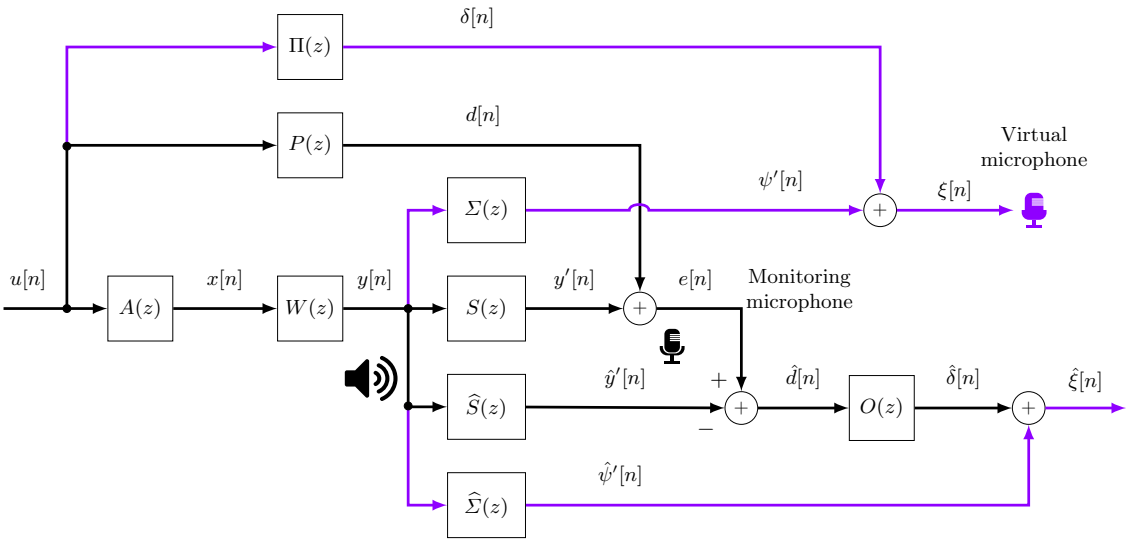


Figure 7.1: Block diagram of a single quiet zone, single reference, SISO fixed ANC system with VMT.

7.1 Fixed Control Filter

A general single quiet zone, single reference, SISO fixed ANC system with VMT is depicted in Fig. 7.1. For the sake of clarity, the arrows directed to the monitoring microphone are drawn in black and the arrows directed to the virtual microphone are depicted in violet.

The unknown signal $u[n]$ is the source disturbance noise to be canceled. This vibration is detected by the accelerometer, placed in a crucial position of the car structure, and filtered by the $A(z)$ block. At the output, reference signal $x[n]$, which can be considered as an informative transformation of the disturbance noise, is obtained.

Signal $u[n]$ propagates into the car cabin and is perceived by both monitoring and virtual error microphone. Filter $P(z)$ is the primary path that represents the physical channel between the vibration and the monitoring microphone. Similarly, it is possible to define the primary path between the disturbance noise and the virtual microphone, i.e., the filter $\Pi(z)$. Their outputs, $d[n]$ and $\delta[n]$, are the respective disturbance signals we wish to control.

Desired signals can be respectively expressed as

$$d[n] = \sum_{q=0}^{Q-1} u[n-q]p[q] = u[n] \otimes p[n]$$
$$\delta[n] = \sum_{q=0}^{Q'-1} u[n-q]\pi[q] = u[n] \otimes \pi[n]$$

where $p[n]$ and $\pi[n]$ are the impulse responses of the filters $P(z)$ and $\Pi(z)$ with filter lengths Q and Q', respectively.

From Fig. 7.1, the block $S(z)$ identifies the acoustic channel between loudspeaker and monitoring microphone (secondary channel), instead, the block $\Sigma(z)$ represents the secondary path associated to the virtual microphone. At the output of these filters, signals $y'[n]$ and $\psi'[n]$ are obtained as discrete convolutions as

$$\begin{aligned} y'[n] &= \sum_{h=0}^{H-1} y[n-h]s[h] \\ &= \sum_{h=0}^{H-1} s[h] \sum_{\ell=0}^{N-1} x[n-\ell-h]w_\ell \\ &= \left(\mathbf{w}^\top \mathbf{x}[n]\right) \otimes s[n] \end{aligned}$$

and

$$\begin{aligned} \psi'[n] &= \sum_{h=0}^{H'-1} y[n-h]\sigma[h] \\ &= \sum_{h=0}^{H'-1} \sigma[h] \sum_{\ell=0}^{N-1} x[n-\ell-h]w_\ell \\ &= \left(\mathbf{w}^\top \mathbf{x}[n]\right) \otimes \sigma[n] \end{aligned}$$

where $s[n]$ and $\sigma[n]$ are the impulse responses of the filters $S(z)$ and $\Sigma(z)$ with lengths H and H', respectively, and N the control filter length.

At the monitoring microphone level, two audio signals are added to each other, namely the disturbance $d[n]$ and the filtered anti-noise $y'[n]$. Therefore, the error

signal is

$$e[n] = d[n] + y'[n] \ . \tag{7.1}$$

Similarly, at the virtual microphone, the error $\xi[n]$ is defined as

$$\xi[n] = \delta[n] + \psi'[n] \ . \tag{7.2}$$

The error signal in (7.2) should be minimized in order to get noise cancellation at the virtual microphone location. However, since $\xi[n]$ is unknown, this error signal can be only estimated. By inverting (7.1), an estimate of the disturbance $\hat{d}[n]$, can be written as

$$\begin{aligned}\hat{d}[n] &= e[n] - \hat{y}'[n] \\ &= e[n] - \sum_{h=0}^{H-1} \hat{s}[h] \sum_{\ell=0}^{N-1} x[n-\ell-h] w_\ell \\ &= e[n] - \left(\mathbf{w}^\top \mathbf{x}[n]\right) \otimes \hat{s}[n]\end{aligned}$$

where $\hat{s}[n]$ is the impulse response of the estimated secondary path between the loudspeaker and monitoring microphone $\widehat{S}(z)$, and $\hat{y}'[n]$ is its output. Once $\hat{d}[n]$ is obtained, it is possible to retrieve the estimate of the disturbance at the virtual microphone, namely $\hat{\delta}[n]$. Mathematically, one gets

$$\begin{aligned}\hat{\delta}[n] &= \sum_{i=0}^{I-1} \hat{d}[n-i] o[i] \\ &= \hat{d}[n] \otimes o[n]\end{aligned}$$

where $o[n]$ is the impulse response of the so-called *observation filter* $O(z)$ of length I which represents the physical channel between the monitoring and virtual microphone. From Fig. 7.1, the anti-noise signal $y[n]$ is filtered by the estimated version of the

secondary path. This yields

$$\begin{aligned}\hat{\psi}'[n] &= \sum_{h=0}^{H'-1} y[n-h]\hat{\sigma}[h] \\ &= \sum_{h=0}^{H'-1} \hat{\sigma}[h] \sum_{\ell=0}^{N-1} x[n-\ell-h]w_\ell \\ &= \left(\mathbf{w}^\top \mathbf{x}[n]\right) \otimes \hat{\sigma}[n]\end{aligned}$$

where $\hat{\sigma}[n]$ is impulse response of the estimated secondary path $\widehat{\Sigma}(z)$ of length H'. Hence, the estimated version of (7.2) becomes

$$\begin{aligned}\hat{\xi}[n] &= \hat{\delta}[n] + \hat{\psi}'[n] \\ &= \hat{\delta}[n] + \left(\mathbf{w}^\top \mathbf{x}[n]\right) \otimes \hat{\sigma}[n] \ . \end{aligned} \tag{7.3}$$

The idea of microphone virtualization is to pursue $\hat{\xi}[n] = \xi[n] = 0$. Several algorithms, e.g., the MMSE algorithm, can be employed in order to achieve this goal. The cost function to be minimized is

$$\chi[n] = \mathbb{E}\left\{\hat{\xi}[n]\hat{\xi}^\top[n]\right\} \ . \tag{7.4}$$

By rewriting (7.3) one gets

$$\begin{aligned}\hat{\xi}[n] &= \hat{\delta}[n] + \hat{\psi}'[n] \\ &= \sum_{i=0}^{I-1} \hat{d}[n-i]o[i] + \sum_{h=0}^{H'-1} \hat{\sigma}[h] \sum_{\ell=0}^{N-1} x[n-\ell-h]w_\ell \\ &= \sum_{i=0}^{I-1} \hat{d}[n-i]o[i] + \sum_{\ell=0}^{N-1} w_\ell \sum_{h=0}^{H'-1} \hat{\sigma}[h]x[n-\ell-h] \\ &= \mathbf{o}^\top \hat{\mathbf{d}}[n] + \sum_{\ell=0}^{N-1} r[n-\ell]w_\ell \\ &= \mathbf{o}^\top \hat{\mathbf{d}}[n] + \mathbf{w}^\top \mathbf{r}[n]\end{aligned} \tag{7.5}$$

where the signal $r[n]$ and vector $\hat{\mathbf{d}}[n]$ are respectively defined as

$$\begin{aligned} r[n] &= \sum_{h=0}^{H'-1} x[n-h]\hat{\sigma}[h] \\ &= \mathbf{x}^\top[n]\hat{\boldsymbol{\sigma}} \\ &= \hat{\boldsymbol{\sigma}}^\top\mathbf{x}[n] \end{aligned}$$

and

$$\hat{\mathbf{d}}[n] = \left[\hat{d}[n], \hat{d}[n-1], ..., \hat{d}[n-I+1]\right]^\top .$$

In particular, the reference signal vector $\mathbf{x}[n]$, the secondary path vector $\hat{\boldsymbol{\sigma}}$ and the vector $\mathbf{r}[n]$ are respectively defined as

$$\begin{aligned} \mathbf{x}[n] &= [x[n], x[n-1], ..., x[n-H'+1]]^\top \\ \hat{\boldsymbol{\sigma}} &= [\sigma[0], \sigma[1], ..., \sigma[H'-1]]^\top \\ \mathbf{r}[n] &= [r[n], r[n-1], ..., r[n-H'+1]]^\top . \end{aligned}$$

Hence, by plugging (7.5) into (7.4), one obtains

$$\begin{aligned} \mathbb{E}\left\{\hat{\xi}[n]\hat{\xi}^\top[n]\right\} &= \mathbb{E}\left\{\left(\mathbf{o}^\top\hat{\mathbf{d}}[n] + \mathbf{w}^\top\mathbf{r}[n]\right)\left(\mathbf{o}^\top\hat{\mathbf{d}}[n] + \mathbf{w}^\top\mathbf{r}[n]\right)^\top\right\} \\ &= \mathbb{E}\left\{\mathbf{o}^\top\hat{\mathbf{d}}[n]\hat{\mathbf{d}}^\top[n]\mathbf{o} + \mathbf{o}^\top\hat{\mathbf{d}}[n]\mathbf{r}^\top[n]\mathbf{w} + \mathbf{w}^\top\mathbf{r}[n]\hat{\mathbf{d}}^\top[n]\mathbf{o} + \mathbf{w}^\top\mathbf{r}[n]\mathbf{r}^\top[n]\mathbf{w}\right\} \\ &= \mathbf{o}^\top\mathbf{C}_{\hat{d}\hat{d}}\mathbf{o} + \mathbf{o}^\top\mathbf{C}_{r\hat{d}}\mathbf{w} + \mathbf{w}^\top\mathbf{C}_{\hat{d}r}\mathbf{o} + \mathbf{w}^\top\mathbf{C}_{rr}\mathbf{w} \end{aligned} \tag{7.6}$$

in which the cross- and auto-correlation matrices are defined as

$$\begin{aligned} \mathbf{C}_{\hat{d}r} &= \mathbb{E}\left\{\mathbf{r}[n]\hat{\mathbf{d}}^\top[n]\right\} \qquad \mathbf{C}_{r\hat{d}} = \mathbb{E}\left\{\hat{\mathbf{d}}[n]\mathbf{r}^\top[n]\right\} \\ \mathbf{C}_{\hat{d}\hat{d}} &= \mathbb{E}\left\{\hat{\mathbf{d}}[n]\hat{\mathbf{d}}^\top[n]\right\} \qquad \mathbf{C}_{rr} = \mathbb{E}\left\{\mathbf{r}[n]\mathbf{r}^\top[n]\right\} \end{aligned}$$

By minimizing (7.6) with respect to the tap-weight vector $\mathbf{w}$ and setting its gradient equal to zero, one gets

$$2\mathbf{C}_{\hat{d}r}\mathbf{o} + 2\mathbf{C}_{rr}\mathbf{w} = 0 \tag{7.7}$$

having exploited the fact that $\mathbf{C}_{\hat{d}r} = \mathbf{C}_{r\hat{d}}^\top$.

The solution of (7.7) is

$$\mathbf{w}^{\text{opt}} = -\mathbf{C}_{rr}^{-1}\mathbf{C}_{\hat{d}r}\mathbf{o} \ .$$

A normalization factor ρ can be introduced in order to prevent the division by zero. Thus, the optimum filter tap-weight vector which minimizes the cost function $\hat{\chi}[n]$ is

$$\mathbf{w}^{\text{opt}} = -\left[\mathbf{C}_{rr} + \rho\mathbf{I}_{H'}\right]^{-1}\mathbf{C}_{\hat{d}r}\mathbf{o} \tag{7.8}$$

where $\mathbf{I}_{H'}$ is the identity matrix of size $H' \times H'$.

Equation (7.8) says that the optimal solution depends on the auto-correlation matrix of vector $\mathbf{r}[n]$ and the cross-correlation between $\hat{d}[n]$ and $\mathbf{r}[n]$. The cross-correlation matrix can be estimated as follows, assuming joint stationarity of all the signals

$$\begin{aligned}
\mathbf{C}_{\hat{d}r} &= \mathbb{E}\left\{\mathbf{r}[n]\hat{\mathbf{d}}^{\top}[n]\right\} \\
&= \mathbb{E}\left\{\begin{bmatrix} r[n] \\ r[n-1] \\ \vdots \\ r[n-H'+1] \end{bmatrix}[\hat{d}[n], \hat{d}[n-1], ..., \hat{d}[n-H'+1]]\right\} \\
&= \mathbb{E}\left\{\begin{matrix} r[n]\hat{d}[n] & \cdots & r[n]\hat{d}[n-H'+1] \\ r[n-1]\hat{d}[n] & \cdots & r[n-1]\hat{d}[n-H'+1] \\ \vdots & \ddots & \vdots \\ r[n-H'+1]\hat{d}[n] & \cdots & r[n-H'+1]\hat{d}[n-H'+1] \end{matrix}\right\} \\
&= \begin{bmatrix} g[0] & g[1] & \cdots & g[H'-1] \\ g[-1] & g[0] & \cdots & g[H'-2] \\ \vdots & \vdots & \ddots & \vdots \\ g[-H'+1] & g[-H'+2] & \cdots & g[0] \end{bmatrix}
\end{aligned}$$

where

$$\begin{aligned}
g[m] &= \mathbb{E}\left\{r[n]\hat{d}[n-m]\right\} \\
&\simeq \frac{1}{N_1 - N_0}\sum_{n=N_0}^{N_1-1} r[n]\hat{d}[n-m]
\end{aligned}$$

with $0 \leq m \leq H' - 1$, $H' < N_0$ and $N_1 \leq N_B$, being N_B the length of the available observation of $r[n]$.

Similarly, the auto-correlation matrix can be estimated as

$$\begin{aligned}\mathbf{C}_{rr} &= \mathbb{E}\left\{\mathbf{r}[n]\mathbf{r}^{\top}[n]\right\} \\ &= \begin{bmatrix} f[0] & f[1] & \cdots & f[H'-1] \\ f[-1] & f[0] & \cdots & f[H'-2] \\ \vdots & \vdots & \ddots & \vdots \\ f[-H'+1] & f[-H'+2] & \cdots & f[0] \end{bmatrix}\end{aligned}$$

where

$$\begin{aligned}f[m] &= \mathbb{E}\left\{r[n]r[n-m]\right\} \\ &\simeq \frac{1}{N_1 - N_0}\sum_{n=N_0}^{N_1-1} r[n]r[n-m] \ .\end{aligned}$$

Finally, an interesting analogy of the virtual microphone ANC system with the standard FF ANC can be made. The virtual microphone ANC system, in fact, differs from the FF one only for the error signal to be minimized. As previously mentioned, the idea is to have $\xi[n] = 0$. From Fig. 7.1, by passing to the $z-$domain, one gets

$$\begin{aligned}\Xi(z) &= \Delta(z) + \Psi'(z) \\ &= U(z)\Pi(z) + Y(z)\Sigma(z) \\ &= U(z)\Pi(z) + X(z)W(z)\Sigma(z) \\ &= U(z)\left[\Pi(z) + A(z)W(z)\Sigma(z)\right] \ .\end{aligned} \tag{7.9}$$

Thus, by equating (7.9) to zero one obtains

$$\Pi(z) + A(z)W(z)\Sigma(z) = 0$$

and its solution is

$$W^{\text{opt}} = -\frac{\Pi(z)}{A(z)\Sigma(z)} \ . \tag{7.10}$$

Neglecting the filter $A(z)$, (7.10) says that the ideal optimal control filter $W^{\text{opt}}(z)$ has

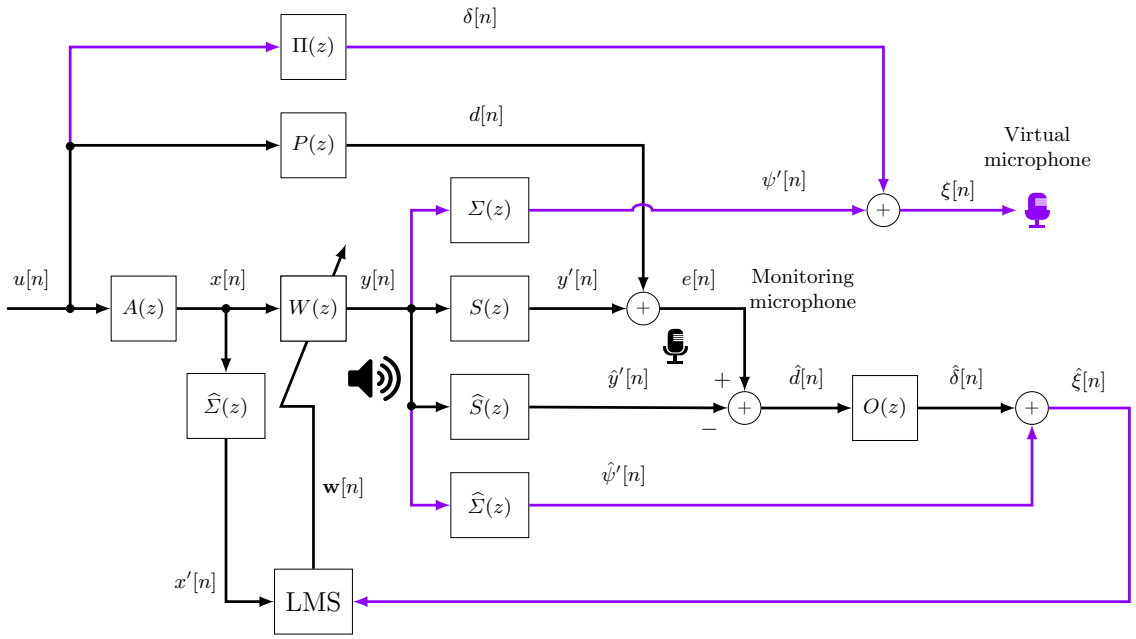


Figure 7.2: Block diagram of a single quiet zone, single reference, SISO adaptive ANC system with VMT.

to be simultaneously proportional to the primary path $\Pi(z)$ and inversely proportional to the secondary path $\Sigma(z)$. By comparing this solution with that of a standard FF ANC system in (A.2), it is possible to conclude that the equations are the same except for the considered primary and secondary path, i.e., $\Pi(z)$ and $\Sigma(z)$ instead of $P(z)$ and $S(z)$, respectively.

7.2 Adaptive Control Filter

A general single quiet zone, single reference SISO adaptive ANC system with VMT is depicted in Fig. 7.2.

In the standard adaptive ANC systems based on the FxLMS algorithm, the cost function to be minimized is the square of the instantaneous error signal $e[n]$. However, the aim of ANC system with VMT is to cancel the disturbance noise in a different position with respect to the monitoring microphone, i.e., that of the virtual one. Thus, in this case, the FxLMS algorithm has to be driven by the estimated error signal at the virtual microphone.

In an adaptive filter approach, the anti-noise $y[n]$ and the time-varying tap-weight vector $\mathbf{w}[n]$ can be written as in (3.1) and (3.3) for $J = K = 1$. The cost function $\chi[n]$ to be minimized with respect to the tap-weight filter $\mathbf{w}[n]$ is the instantaneous squared error

$$\chi[n] = \hat{\xi}^2[n] \ .$$

Similarly to the derivation in Section 3.1, its gradient can be calculated as follows

$$\begin{aligned}
\frac{\partial \chi[n]}{\partial w_i} &= \frac{\hat{\xi}^2[n]}{\partial w_i} \\
&= 2\hat{\xi}[n]\frac{\hat{\xi}[n]}{\partial w_i} \\
&= 2\hat{\xi}[n]\frac{\partial}{\partial w_i}\left(\hat{\delta}[n] + \hat{\psi}'[n]\right) \\
&= 2\hat{\xi}[n]\left[\hat{\sigma}[n] \otimes x[n-i]\right] \\
&= 2\hat{\xi}[n]x'[n-i]
\end{aligned}$$

where $x'[n-i] = \hat{\sigma}[n] \otimes x[n-i]$ for $i = 1, 2, ..., N-1$.

By defining

$$\begin{aligned}
\mathbf{x}'[n] &= \left[x'[n], x'[n-1], ..., x'[n-N+1]\right]^\top \\
&= \left[\hat{\sigma}[n] \otimes x[n], \hat{\sigma}[n] \otimes x[n-1], ..., \hat{\sigma}[n] \otimes x[n-N+1]\right]^\top
\end{aligned}$$

the tap-weight recursion of the leaky normalized FxLMS algorithm for a SISO ANC system with VMT becomes

$$\boxed{\mathbf{w}[n+1] = \lambda\mathbf{w}[n] + \tilde{\mu}\frac{\mathbf{x}'[n]}{\alpha + \mathbf{x}'^\top[n]\mathbf{x}'[n]}\hat{\xi}[n] \ .}$$

Chapter 8

Observation Filter Estimation

In this chapter, the main concepts on the observation filter estimation are discussed. In Section 8.1, the filter causality issue is presented. Then, in Section 8.2, the main mathematical background of the adopted observation filter estimation algorithms, by means multiple monitoring microphones for multiple quiet zones are derived. More precisely, the observation filters are estimated for the LMS, Recursive Least Squares (RLS) and MMSE algorithm, in Sections 8.2.1, 8.2.2 and 8.2.3, respectively. Finally, in Section 8.3, the problem on the robustness of the observation filter under virtual microphone mismatch and road scenario mismatch are discussed.

8.1 Causality Issue

As mentioned in Chapter 7, the aim of the observation filter is to retrieve the signal $\hat{\delta}[n]$ (unknown) from the signal $\hat{d}[n]$ (estimate of $d[n]$). In order to analyze the behavior of the observation filter it may be convenient to operate in the z-domain.

At the virtual microphone level, the error signal $\xi[n]$ can be expressed as in (7.9), repeated here for convenience

$$\begin{aligned} \Xi(z) &= \Delta(z) + \Psi'(z) \\ &= U(z)\left[\Pi(z) + A(z)W(z)\Sigma(z)\right] \ . \end{aligned} \tag{8.1}$$

Similarly, the estimated error signal $\hat{\xi}[n]$ becomes

$$\begin{aligned}\widehat{\Xi}(z) &= \widehat{\Delta}(z) + \widehat{\Psi}'(z) \\ &= \widehat{D}(z)O(z) + Y(z)\widehat{\Sigma}(z) \\ &= [E(z) - Y'(z)]\,O(z) + X(z)W(z)\widehat{\Sigma}(z) \\ &= \left\{U(z)\left[P(z) + A(z)W(z)\left(S(z) - \widehat{S}(z)\right)\right]\right\}O(z) + U(z)A(z)W(z)\widehat{\Sigma}(z)\ . \end{aligned} \tag{8.2}$$

Under the assumption of perfect estimation of secondary paths, i.e., $\widehat{S}(z) = S(z)$ and $\widehat{\Sigma}(z) = \Sigma(z)$, (8.2) reduces to

$$\widehat{\Xi}(z) = U(z)\left[P(z)O(z) + A(z)W(z)\Sigma(z)\right]\ . \tag{8.3}$$

Hence, by subtracting (8.3) from (8.1), and setting the result equal to zero, one gets

$$\Xi(z) - \widehat{\Xi}(z) = U(z)\left[\Pi(z) - P(z)O(z)\right] = 0\ . \tag{8.4}$$

The solution of (8.4) is thus

$$\boxed{O^{\text{opt}}(z) = \frac{\Pi(z)}{P(z)}\ .} \tag{8.5}$$

Equation (8.5) says that the observation filter has to be proportional to the primary path $\Pi(z)$ and inversely proportional to the $P(z)$. Unfortunately, since the primary paths are unknown, (8.5) can only be considered as a theoretical solution to which an observation filter has to tend in order to obtain good estimation performance.

From a computational viewpoint, an issue of the theoretical solution (8.5) is represented by the inversion of the primary path $P(z)$. At the same time, a problem of system causality may cause significant performance degradation. Let us suppose that the primary path $\Pi(z)$ introduces a delay larger than the primary path $P(z)$. This means that the disturbance noise is first perceived by the monitoring microphone and secondly by the virtual one. This yields a causal observation filter $O(z)$ and the computational complexity is exclusively due to the inversion of the primary path $P(z)$. On the contrary, if the disturbance noise is first perceived by the virtual microphone, e.g., due to the car cabin geometry, in order to compensate for this delay the observation filter should anticipate the signal, which is obviously physically impossible.

In order to cope with this limitation, a delayed observation filter $O^{\mathrm{D}}(z)$ can be introduced such that its output can be written as

$$\sum_{i=0}^{I-1} d[n-i]o^{\mathrm{D}}[i] = \hat{\delta}[n-n_0]$$

being n_0 the delay in samples. In fact, this allows to retrieve the signal $\hat{\delta}[n-n_0]$, i.e., $\hat{\delta}[n]$ but for a delay n_0. The delayed observation filter $O^{\mathrm{D}}(z)$ can be estimated by adopting both fixed and adaptive approaches as will be discussed in Section 8.2.

Since an introduction of a too high delay may cause significant system performance degradation, to find a good value of the delay n_0 may not be trivial. One approach may be the evaluation of distance between monitoring and virtual microphone, e.g., if the microphone-to-microphone distance is 15 cm, and a sampling frequency of $f_s = 48$ Hz is used, a reasonable delay is $n_0 = 21$ samples by using the speed of sound $v_0 = 343.8$ m/s in air at 20°C. However, since a sound propagating inside the cabin of a car is usually reflected several times, this solution may not be sufficient.

An approach based on trial and error can be considered in order to account for possible significant non-causal components of the observation filter impulse response to be included [53].

Similar to the microphone-to-microphone distance evaluation method, another approach is the estimation of the delay spread of the primary paths by analyzing their impulse responses. Although this measure is unconventional and imprecise, hereinafter it will be used in order to identify a causal and non-causal system. Explicitly, if the peak of the impulse response of $P(z)$ arrives before that the maximum peak of $\Pi(z)$, the system is considered "causal". On the contrary, if the maximum peak of the impulse response of $P(z)$ arrives after that the maximum peak of $\Pi(z)$, the system is considered "non-causal". A simplified example of causal and non-causal systems by observing the impulse responses of the primary paths is depicted in Figs. 8.1(a) and (b), respectively.

8.2 Considered Algorithms

In this section, the considered algorithms for the OFs estimation are presented.

Similarly to what occurs for the estimation of secondary paths, a preliminary iden-

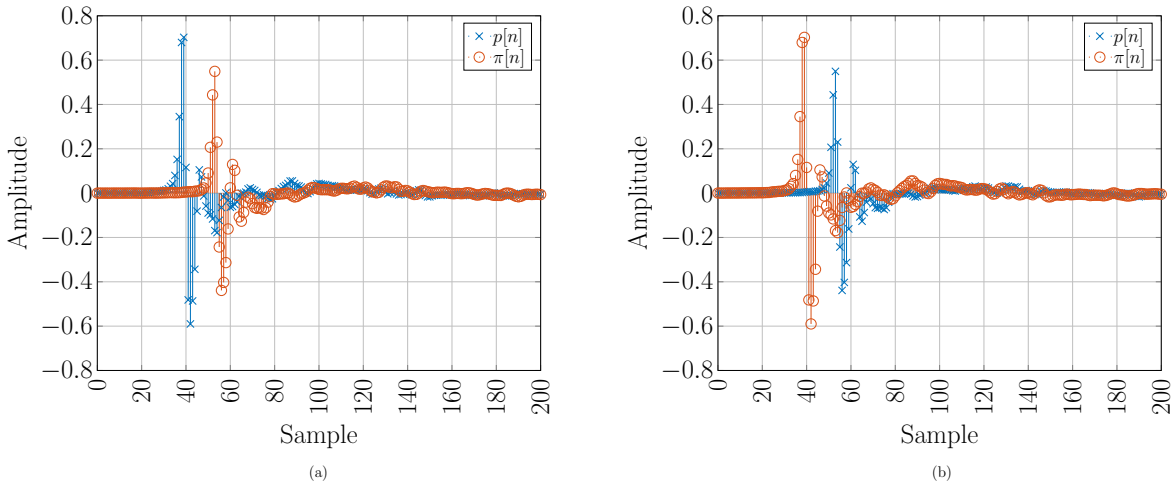


Figure 8.1: Simplified example on determining the causality of the system: (a) causal system, (b) non-causal system.

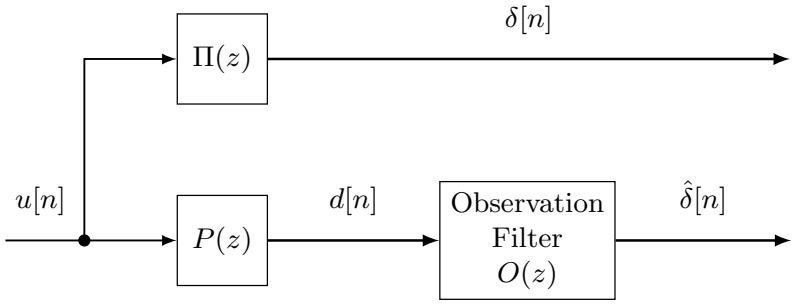


Figure 8.2: General block diagram of environmental audio signals acquired by single monitoring and single virtual microphone and reconstruction by observation filter.

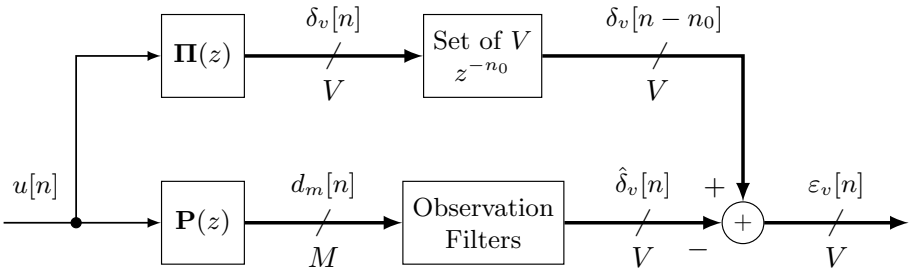


Figure 8.3: General block diagram of environmental audio acquired by monitoring and virtual microphones and their reconstructions by delayed observation filters .

tification stage in which a physical microphone is momentarily placed at the virtual microphone position is required for observation filter estimation purposes. This estimation is performed "off-line", i.e., when the ANC system is OFF. This yields the anti-noise signal equal to zero, i.e., $y[n] = 0$. Thus, the block diagram depicted in Figs. 7.1 and 7.2 reduces to Fig. 8.2. From this viewpoint, it is trivial to observe that the observation filter $O(z)$ must tend to (8.5) in order to have $\hat{\delta}[n] = \delta[n]$.

When multiple virtual and monitoring microphones are considered, the block diagram becomes 8.3. Note that, for the sake of simplicity, only the scenario in which multiple microphones are employed in order to virtualize several quiet zones is considered. Moreover, due to particular geometrical car cabin characteristics, a delay of n_0 samples, introduced in the sequence of the virtual microphone signals, is always required. For this reason, the notation for the delayed observation filter $O^{\mathrm{D}}(z)$ it is simply substituted by $O(z)$.

From Fig. 8.3, the unknown environmental audio signal $u[n]$ propagating within the car cabin is acquired by M monitoring and V virtual microphones. The physical acoustic channels between the vibration source and the microphones are referred to as primary paths and modeled by FIR filters with transfer functions $\mathbf{P}(z)$ and $\mathbf{\Pi}(z)$ as mentioned in Chapter 7. Hence, the m-th monitoring and v-th virtual microphone

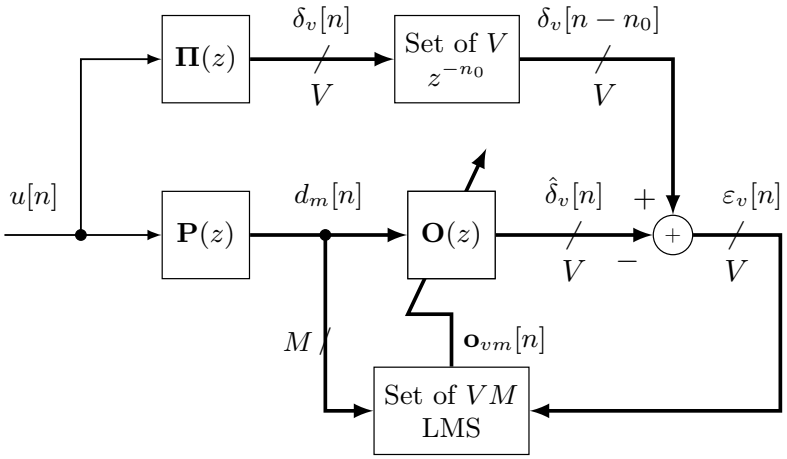


Figure 8.4: Block diagram for the reconstruction of virtual microphone signals by observation filters estimated by means of the LMS algorithm.

signals can be respectively expressed as

$$d_m[n] = \sum_{q=0}^{Q-1} u[n-q]p_m[q] = u[n] \otimes p_m[n] \tag{8.6}$$

$$\delta_v[n] = \sum_{q'=0}^{Q'-1} u[n-q']\pi_v[q'] = u[n] \otimes \pi_v[n] \tag{8.7}$$

where $p_m[n]$ and $\pi_v[n]$ are the m-th and v-th impulse responses associated to $\mathbf{P}(z)$ and $\mathbf{\Pi}(z)$ with lengths Q and Q', respectively.

The aim of the observation filters is to retrieve the V virtual microphone signals from the M monitoring ones. Thus, at the output of the observation filter, the reconstructed version of the v-th virtual microphone $\hat{\delta}_v[n]$ is obtained. A set of V delays of n_0 samples is introduced in the set of the virtual microphone signals to account for the causality issue previously discussed. Thus, at the v-th virtual microphone, signal $\delta_v[n-n_0]$ is obtained as shown in Fig. 8.3. Finally, the v-th error signal can be expressed as

$$\varepsilon_v[n] = \delta_v[n-n_0] - \hat{\delta}_v[n] \ . \tag{8.8}$$

8.2.1 Least Mean Square (LMS)

The LMS block diagram for observation filter estimation is depicted in Fig. 8.4. The disturbing signal $u[n]$ feeds the primary paths $\mathbf{P}(z)$ and $\mathbf{\Pi}(z)$, yielding the signals $d_m[n]$ and $\delta_v[n]$, for the m-th monitoring and v-th virtual microphones as expressed in (8.6) and (8.7). Then, the v-th error signal can be written as (8.8).

In particular, the retrieved version of the v-th virtual microphone signal is

$$\begin{aligned}\hat{\delta}_v[n] =& \sum_{m=0}^{M-1}\left(\sum_{i=0}^{I-1} o_{vmi}[n]d_m[n-i]\right)\\ =& \sum_{m=0}^{M-1}\mathbf{d}_m^\top[n]\mathbf{o}_{vm}[n]\\ =& \sum_{m=0}^{M-1}\mathbf{o}_{vm}^\top[n]\mathbf{d}_m[n]\end{aligned}$$

where the input vector is defined as

$$\mathbf{d}_m[n] = [d_m[n], d_m[n-1], ..., d_m[n-I+1]]^\top \tag{8.9}$$

where I is the observation filter length and

$$\mathbf{o}_{vm}[n] = [o_{vm0}[n], o_{vm1}[n], ..., o_{vm,I-1}[n]]^\top \tag{8.10}$$

is the impulse response vector at the n-th time epoch of the adaptive filter from the v-th virtual microphone signal to the m-th monitoring one. In the LMS algorithm, the cost function is approximated by the instantaneous estimate of the squared error signal. Then, for the v-th virtual microphone, one gets

$$\nabla\xi_v[n] = \nabla\varepsilon_v^2[n] \tag{8.11}$$

and the ℓ-th element of the stochastic gradient of this cost function is evaluated as

follows

$$\begin{aligned}\frac{\partial \varepsilon_v^2[n]}{\partial o_{vm\ell}} &= 2\varepsilon_v[n]\frac{\partial \varepsilon_v[n]}{\partial o_{vm\ell}} \\ &= 2\varepsilon_v[n]\frac{\partial}{\partial o_{vm\ell}}\left(\delta_v[n-n_0]-\hat{\delta}_v[n]\right) \\ &= -2\varepsilon_v[n]\frac{\partial}{\partial o_{vm\ell}}\left[\sum_{m'=1}^{M}\left(\sum_{\ell'=0}^{I-1} d_{m'}[n-\ell']o_{v'm'\ell'}\right)\right] \\ &= -2\varepsilon_v[n]d_m[n-\ell]\end{aligned}$$

in which the independence between $o_{vm\ell}$ and $\delta_v[n]$ and the fact that the partial derivative of the double sum is different from zero if and only if $v'=v$, $m'=m$ and $\ell'=\ell$ were exploited. Thus, in vectorial form it is possible to compactly write (8.11) as

$$\nabla\varepsilon_v^2[n] = -2\varepsilon_v[n]\mathbf{d}_m[n] \ .$$

Finally, the tap-weight update equation of the leaky normalized LMS algorithm for the observation filter associated to the m-th monitoring and v-th virtual microphone at the n-th time epoch is obtained. Adopting a leaky normalized LMS as discussed in Section 1.3, the recursion is

$$\mathbf{o}_{vm}[n+1] = \lambda\mathbf{o}_{vm}[n] + \mu\frac{\mathbf{d}_m[n]}{\alpha+\mathbf{d}_m^\top[n]\mathbf{d}_m[n]}\varepsilon_v[n] \ . \tag{8.12}$$

8.2.2 Recursive Least Squares (RLS)

Differently from the LMS algorithm, presented in Section 8.2.1, which minimizes the mean square error, the aim of RLS algorithm is to recursively find the coefficients which minimize the linear least squares cost function related to the input signals [17]. The corresponding block diagram is the same of what was shown in Fig. 8.4 where, in this case the block "Set of VM RLS" is present instead of the LMS one.

The most significant advantage exhibited by this algorithm is its extremely fast convergence at the expense of a high computational complexity.

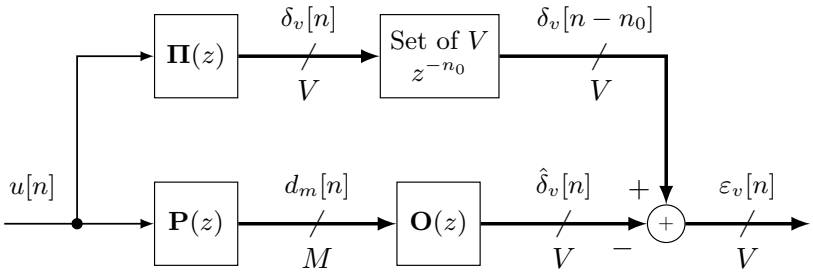


Figure 8.5: Block diagram for the reconstruction of virtual microphone signals by observation filters estimated by means of the MMSE algorithm.

The RLS algorithm is implemented by the following mathematical steps:

1. $\mathbf{k}[n] = \frac{\beta^{-1}\mathbf{B}[n-1]\mathbf{u}[n]}{1+\beta^{-1}\mathbf{u}^\top[n]\mathbf{B}[n-1]\mathbf{u}[n]}$
2. $\hat{\delta}_v[n] = \mathbf{o}_{vm}\top[n-1]\mathbf{u}[n]$
3. $\varepsilon_v[n] = \delta_v[n-n_0] - \hat{\delta}_v[n]$
4. $\mathbf{o}_{vm}[n] = \mathbf{o}_{vm}[n-1] + \mathbf{k}^\top[n]\varepsilon_v[n]$
5. $\mathbf{B}[n] = \beta^{-1}\mathbf{B}[n-1] - \beta^{-1}\mathbf{k}[n]\mathbf{u}^\top[n]\mathbf{B}[n-1]$

where β is the reciprocal of the forgetting factor, $\mathbf{k}[n]$ is the gain vector at step n and $\mathbf{B}[n]$ is the inverse covariance matrix at step n.

8.2.3 Minimum Mean Square Error (MMSE)

The block diagram for the estimation of the observation filters by means of the MMSE algorithm is depicted in Fig. 8.5. The reconstructed version of the v-th virtual microphone signal can be expressed as

$$\begin{aligned}\hat{\delta}_v[n] &= \mathbf{d'}^\top[n]\mathbf{o}'_v \\ &= \mathbf{o}'^\top_v\mathbf{d'}[n]\end{aligned}$$

where the vectors of the monitoring microphone signals and impulse responses of the observation filters are defined as

$$\begin{aligned}\mathbf{d}'[n] &= \left[\mathbf{d}_1^\top[n], \mathbf{d}_2^\top[n], ..., \mathbf{d}_m^\top[n], ..., \mathbf{d}_M^\top[n]\right]^\top \\ \mathbf{o}'_v &= \left[\mathbf{o}_{v1}^\top, \mathbf{o}_{v2}^\top, ..., \mathbf{o}_{vm}^\top, ..., \mathbf{o}_{vM}^\top\right]^\top\end{aligned} \tag{8.13}$$

in which the vectors $\mathbf{d}_m[n]$ and $\mathbf{o}_{vm}$ were previously defined in (8.9) and (8.10). Note that, since the vector of the impulse responses is now time-independent, the index n in (8.13) has been dropped.

The mean square value of the v-th error signal $\varepsilon_v[n]$ can be evaluated as follows

$$\begin{aligned}\mathbb{E}\left\{\varepsilon_v[n]\varepsilon_v^\top[n]\right\} &= \mathbb{E}\left\{\left(\delta_v[n-n_0] - \hat{\delta}_v[n]\right)\left(\delta_v[n-n_0] - \hat{\delta}_v[n]\right)^\top\right\} \\ &= \mathbb{E}\left\{\left(\delta_v[n-n_0] - \mathbf{o}_v'^\top\mathbf{d}'[n]\right)\left(\delta_v[n-n_0] - \mathbf{o}_v'^\top\mathbf{d}'[n]\right)^\top\right\} \\ &= c_{\delta_v} - \mathbf{c}_{\mathbf{d}'\delta_v}^\top\mathbf{o}'_v - \mathbf{o}_v'^\top\mathbf{c}_{\mathbf{d}'\delta_v} + \mathbf{o}_v'^\top\mathbf{C}_{\mathbf{d}'\mathbf{d}'}\mathbf{o}'_v \\ &= c_{\delta_v} - 2\mathbf{c}_{\mathbf{d}'\delta_v}^\top\mathbf{o}'_v + \mathbf{o}_v'^\top\mathbf{C}_{\mathbf{d}'\mathbf{d}'}\mathbf{o}'_v\end{aligned} \tag{8.14}$$

where $c_{\delta_v} = \mathbb{E}\left\{\delta_v^2[n-n_0]\right\}$ denotes the mean square value of the v-th virtual microphone signal, $\mathbf{c}_{\mathbf{d}'\delta_v} = \mathbb{E}\left\{\mathbf{d}'[n]\delta_v[n-n_0]\right\}$ identifies the cross-correlation vector between the monitoring microphone vector and the v-th virtual microphone signal and $\mathbf{C}_{\mathbf{d}'\mathbf{d}'} = \mathbb{E}\left\{\mathbf{d}'[n]\mathbf{d}'^\top[n]\right\}$ represents the auto-correlation matrix of monitoring microphone vector. By minimizing (8.14) with respect to the v-th observation filter tap-weights by equating its gradient to zero, one gets

$$\mathbf{o}'_v = \left[\mathbf{C}_{\mathbf{d}'\mathbf{d}'} + \rho\mathbf{I}_{MI}\right]^{-1}\mathbf{c}_{\mathbf{d}'\delta_v} \tag{8.15}$$

where ρ is the regularization factor introduced in (7.8) since the auto-correlation matrix may be ill-conditioned [53, 54] and $\mathbf{I}_{MI}$ denotes the identity matrix of size $MI \times MI$.

The cross-correlation vector $\mathbf{c}_{\mathbf{d}'\delta_v}$ can be expressed as

$$\begin{aligned}\mathbf{c}_{\mathbf{d}'\delta_v} &= \mathbb{E}\left\{\begin{bmatrix}\mathbf{d}_1[n]\\ \mathbf{d}_2[n]\\ \vdots\\ \mathbf{d}_M[n]\end{bmatrix}\delta_v[n-n_0]\right\}\\ &= [g_{v1}[0], g_{v1}[1], ..., g_{v1}[I-1], ..., g_{vM}[0], g_{vM}[1], ..., g_{vM}[I-1]]^\top . \end{aligned} \quad (8.16)$$

Similarly to [55], the element $g_{vm}[t]$ can be estimated as a temporal correlation on a window of length $N_1 - N_0$, i.e.,

$$\begin{aligned} g_{vm}[t] &= \mathbb{E}\left\{d_m[n-t]\delta_v[n-n_0]\right\}\\ &\simeq \frac{1}{N_1-N_0}\sum_{n=N_0}^{N_1-1} d_m[n-t]\delta_v[n-n_0] \end{aligned} \quad (8.17)$$

where $0 \leq t \leq I-1 \leq N_0$ for $v = 1, 2, ..., V$ and $m = 1, 2, .., M$. The auto-correlation matrix $\mathbf{C}_{\mathbf{d}'\mathbf{d}'}$ is defined as

$$\begin{aligned}\mathbf{C}_{\mathbf{d}'\mathbf{d}'} &= \mathbb{E}\left\{\mathbf{d}'[n]\mathbf{d}'^\top[n]\right\}\\ &= \mathbf{E}\left\{\begin{bmatrix}\mathbf{d}_1[n]\mathbf{d}_1^\top[n] & \mathbf{d}_1[n]\mathbf{d}_2^\top[n] & \cdots & \mathbf{d}_1[n]\mathbf{d}_M^\top[n]\\ \mathbf{d}_2[n]\mathbf{d}_1^\top[n] & \mathbf{d}_2[n]\mathbf{d}_2^\top[n] & \cdots & \mathbf{d}_2[n]\mathbf{d}_M^\top[n]\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{d}_M[n]\mathbf{d}_1^\top[n] & \mathbf{d}_M[n]\mathbf{d}_2^\top[n] & \cdots & \mathbf{d}_M[n]\mathbf{d}_M^\top[n]\end{bmatrix}\right\} . \end{aligned} \quad (8.18)$$

From (8.18), it is possible to define the square matrix $\mathbf{R}_{\ell k}$ of size $I \times I$ as follows

$$\begin{aligned}\mathbf{R}_{\ell k} &= \mathbb{E}\left\{\mathbf{d}_\ell[n]\mathbf{d}_k^\top[n]\right\} \\ &= \mathbb{E}\left\{\begin{bmatrix} d_\ell[n] \\ d_\ell[n-1] \\ \vdots \\ d_\ell[n-I+1] \end{bmatrix}[d_k[n], d_k[n-1], ..., d_k[n-I+1]]\right\} \\ &= \begin{bmatrix} r_{lk}[0] & r_{lk}[1] & \cdots & r_{lk}[I-1] \\ r_{lk}[-1] & r_{lk}[0] & \cdots & r_{lk}[I-2] \\ \vdots & \vdots & \ddots & \vdots \\ r_{lk}[-I+1] & r_{lk}[-I+2] & \cdots & r_{lk}[0] \end{bmatrix} . \end{aligned} \tag{8.19}$$

Similarly to (8.17), the element $r_{\ell k}[t]$ can be estimated as a temporal correlation over the window length $N_1 - N_0$ such as

$$\begin{aligned} r_{\ell k}[t] &= \mathbb{E}\left\{d_\ell[n]d_k[n-t]\right\} \\ &\simeq \frac{1}{N_1 - N_0}\sum_{n=N_0}^{N_1-1} d_\ell[n]d_k[n-t] \end{aligned}$$

in which $0 \leq t \leq I-1 \leq N_0$ for $r, l = 1, 2, ..., M$. In particular, the following symmetry property between the signals $d_k[n]$ and $d_\ell[n]$ is valid under the assumption of cross stationarity

$$\begin{aligned} r_{k\ell}[t] &= \mathbb{E}\left\{d_k[n]d_\ell[n-t]\right\} \\ &= \mathbb{E}\left\{d_k[n+t]d_\ell[n-t+t]\right\} \\ &= \mathbb{E}\left\{d_\ell[n]d_k[n+t]\right\} \\ &= r_{\ell k}[-t] . \end{aligned} \tag{8.20}$$

Thus, the matrix $\mathbf{R}_{\ell k}$ in (8.19) can be written as

$$\mathbf{R}_{\ell k} = \begin{bmatrix} r_{\ell k}[0] & r_{\ell k}[1] & \cdots & r_{\ell k}[I-1] \\ r_{k\ell}[1] & r_{\ell k}[0] & \cdots & r_{\ell k}[I-1] \\ \vdots & \vdots & \ddots & \vdots \\ r_{k\ell}[I-1] & r_{k\ell}[I-2] & \cdots & r_{\ell k}[0] \end{bmatrix}. \tag{8.21}$$

Finally, by exploiting the property in (8.20) and collecting the definitions in (8.21) and (8.18), the auto-correlation matrix $\mathbf{C}_{\mathbf{d}'\mathbf{d}'}$ in (8.15) can be written as

$$\begin{aligned} \mathbf{C}_{\mathbf{d}'\mathbf{d}'} &= \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \cdots & \mathbf{R}_{1M} \\ \mathbf{R}_{21} & \mathbf{R}_{22} & \cdots & \mathbf{R}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{M1} & \mathbf{R}_{M2} & \cdots & \mathbf{R}_{MM} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \cdots & \mathbf{R}_{1M} \\ \mathbf{R}_{12}^{\top} & \mathbf{R}_{22} & \cdots & \mathbf{R}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{1M}^{\top} & \mathbf{R}_{2M}^{\top} & \cdots & \mathbf{R}_{MM} \end{bmatrix}. \end{aligned} \tag{8.22}$$

From a practical viewpoint, the evaluation of (8.22) can be implemented by evaluating only the diagonal matrices and the upper (or lower) triangular ones. In fact, considering for example $M = 4$, the following structure arises

$$\mathbf{C}_{\mathbf{d}'\mathbf{d}'} = \begin{bmatrix} \mathbf{A} & \mathbf{E} & \mathbf{F} & \mathbf{G} \\ \mathbf{E}^{\top} & \mathbf{B} & \mathbf{H} & \mathbf{I} \\ \mathbf{F}^{\top} & \mathbf{H}^{\top} & \mathbf{C} & \mathbf{J} \\ \mathbf{G}^{\top} & \mathbf{I}^{\top} & \mathbf{J}^{\top} & \mathbf{D} \end{bmatrix}$$

where the blocks are solely indicative of the matrix structure.

8.3 Robustness Testing

In the previous sections, the issues due to the observation filter estimate were analyzed. In the following section, robustness of estimated observation filter under virtual

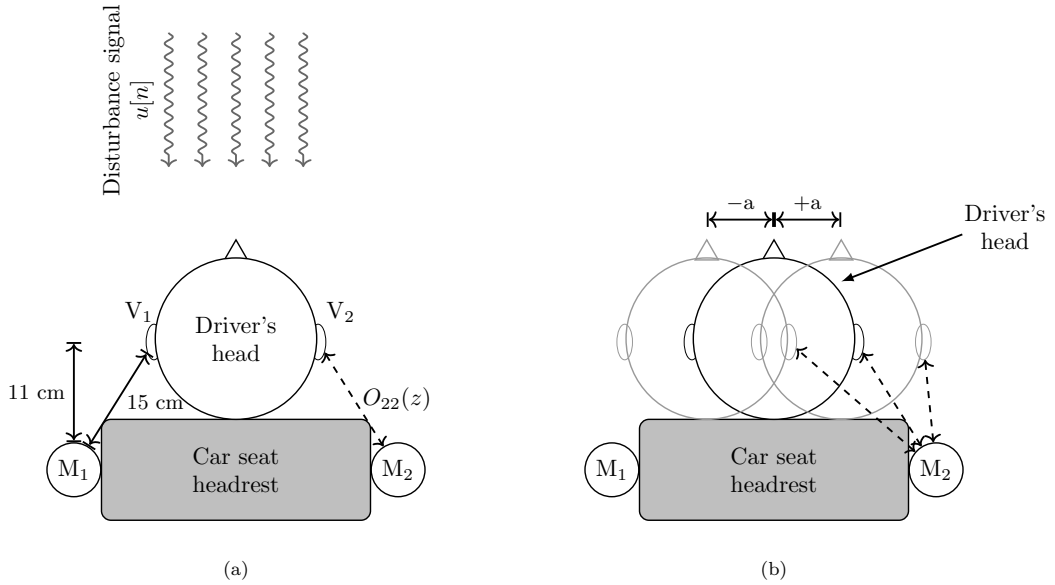


Figure 8.6: Geometrical setup of a car seat headrest: (a) top view, (b) head shift from top view.

microphone and road scenario mismatch is addressed.

As mentioned in Section 8.2, observation filter estimation is performed off-line, i.e., when the ANC system is OFF. This means that, during this estimation phase, signals $d_v[n]$ and $\delta_v[n]$ are acquired by physical microphones placed at monitoring and virtual locations, respectively. A quiet zone around the driver's ears is often preferable, e.g., by employing a manikin equipped with two binaural microphones.

Let us consider a car seat headrest as depicted in Fig. 8.6(a) equipped with a monitoring microphone, namely M_2 and a binaural microphone (employed as a virtual) placed at the right manikin's ear, namely V_2. This setup yields a single quiet zone SISO ANC system with monitoring and virtual microphones M_2 and V_2, respectively. Once the observation filter impulse response estimation is performed, e.g., by adopting one of the approaches described in Section 8.2, this impulse response is implemented in block $\mathbf{O}(z)$ of Fig. 7.2 remaining stable and immutable. Let us suppose now that, while the car is running and the ANC system is ON, the driver's head has a shift of $\pm a$ from the initial position as depicted in Fig. 8.6(b). Due to the head shift, the channel between monitoring and virtual microphones is changed. Is the observation filter $\mathbf{O}(z)$ previously estimated sufficiently robust to this shift? How much is the maximum head

shift that can be dealt with the by the system with negligible performance degradation? This is a key analysis if the distance between monitoring and virtual microphones becomes larger, e.g., when the monitoring microphone is not positioned by the headrest.

Finally, in order to test the observation filters estimation accuracy, a robustness test under road mismatch is considered. More precisely, the observation filters are estimated in a particular driving/road scenario, e.g., constant or variable speed, rough or smooth asphalt, and then they are employed in a different one.

Chapter 9

Experimental Results

Results on estimation of the observation filter are presented in this chapter. In Section 9.1, system performance indicators are defined. Numerical results are organized according to the acquisition method of monitoring microphone signal $d[n]$ and virtual microphone signal $\delta[n]$. In particular, we distinguish between "indirect" and "direct" measurements.

In the indirect measurement procedure, presented in Section 9.2, since the virtual microphone signal $\delta[n]$ is not available, it is obtained by filtering an input signal (realistic measurement or white gaussian noise) through an FIR filter, that can be both synthetic (see Section 9.2.1) and experimental (see Section 9.2.2). The observation filter estimation is performed by applying the MMSE and LMS approach.

On the contrary, in the direct measurement procedure, the virtual microphone signal $\delta[n]$ is available, i.e., by previous signal acquisitions based on employing microphones positioned in the desired region. Section 9.3 is dedicated to this case.

In particular, numerical results are organized based on the considered measurement setup, i.e., sedan (class D), luxury car (class S) in Sections 9.3.1 and 9.3.2, respectively. Finally, in Section 9.3.3, results on two experimental setups are presented. For all the considered cases, numerical results on the robustness of the observation filter estimation under road mismatch are shown in the corresponding section. Note that, both the virtualization of a single quiet zone performed by a single monitoring microphone and multiple quiet zones by means multiple monitoring microphones are considered.

9.1 Performance Indicators

The obtained results are assessed in terms of the following performance measures:

- Average SPL of virtual signal $\delta[n]$, its reconstructed version $\hat{\delta}[n]$ and the corresponding error $\varepsilon[n]$.
- Spectral coherence $C_{\mathbf{D},\delta}$ between monitoring microphones and the virtual ones.
- Sliding window SPL $\mathrm{L}_p[n]$ of disturbance signal $\delta[n]$ and error signal $\xi[n]$ (only when ANC system is ON).
- Mean Square Error (MSE) Υ evaluated on error signal $\varepsilon[n]$ normalized with respect to the input one $\delta[n]$.

Average SPL and spectral coherence measures were previously defined in Section 2.2. More precisely, in order to investigate the physical limitations of the considered empirical setups, a preliminary measurement on the spectral coherence between monitoring and virtual microphones can be pursued. The higher the spectral correlation between such signals, the easier the observation filter task, the better the performance on the estimation accuracy. For the v-th virtual microphone signal $\delta_v[n]$ and a matrix of monitoring microphone signals $\mathbf{D} = [\mathbf{d}_1[n], \mathbf{d}_2[n], ..., \mathbf{d}_M[n]]^\top$, the multiple spectral coherence $\mathbf{C}_{\mathbf{D},\delta_v}(f)$ is defined as in (2.2). The sliding window SPL (in dB) is defined in (6.1).

The accuracy of the observation filter estimation is assessed in terms of MSE normalized with respect to the mean square value of the input signal. For the v-th virtual microphone, the normalized MSE, in dB scale, is defined as

$$\begin{aligned}\Upsilon_v &= 10\log_{10}\frac{\sum_{n=0}^{N-1}\varepsilon_v^2[n]}{\sum_{n=0}^{N-1}\delta_v^2[n-n_0]} \\ &= 10\log_{10}\frac{\sum_{n=0}^{N-1}\left(\delta_v[n-n_0]-\hat{\delta}_v[n]\right)^2}{\sum_{n=0}^{N-1}\delta_v^2[n-n_0]} \quad [\mathrm{dB}]\end{aligned} \tag{9.1}$$

where N indicates the time window length (in samples) for MSE evaluation. Ideally, when the reconstructed signal $\hat{\delta}_v[n]$ equals the target one $\delta_v[n - n_0]$, the error signal $\varepsilon_v[n]$ tends to zero and the MSE $\Upsilon_v \to -\infty$ consequently. Moreover, to have an efficient performance indicator over the whole frequency band, the MSE may be evaluated for 1- and 1/3-octave bands. This means that the error and input signal in (9.1) are decomposed into octave and fractional-octave sub-bands. In this context, it is important to say that the observation filter parameters, i.e., the length I, the delay n_0, the step-size μ, leakage factor λ in (8.12) and the regularization factor ρ in (8.15) are empirically set in order to maximize the OF estimation accuracy in terms of MSE.

9.2 Indirect Measurements

In this section, the results of the so-called indirect measurements are presented. In this case, since the virtual microphone signals are not available, they are synthesized starting from the monitoring ones. This scenario can be considered as a preliminary stage in which we have to cope with the lack of such realistic measurements and for these reasons the main most important results are presented only.

9.2.1 Synthetic Response

An analysis of observation filter estimation performance with the use of synthesized primary path filters is presented here. These numerical results can be considered as a starting point in order to move to a more realistic and complicated setup with experimental primary path filters.

9.2.1.1 Fixed Approach: MMSE

The considered system is a single quiet zone, based on multiple references as depicted in Fig. 7.2.Hence, results on ANC system with VMT are presented. The OF estimation is performed by employing the fixed approach, i.e., the MMSE method, as described in Section 8.2.3.

The input signal $d[n]$ is provided by an experimental measurement campaign performed by ASK Industries on a realistic car interior, i.e., that of the sedan, presented in Section 2.1. Two binaural microphones, placed by the diver's left and right ears,

Table 9.1: Table on employed microphone signal (causal system).

	$\delta[n]$
`CASE 1`	$d[n-1]$
`CASE 2`	$d[n] \otimes \varphi[n]$

Table 9.2: Summary table on employed microphone signal (non-causal system).

	$\delta[n]$
`CASE 1`	$d[n]$
`CASE 2`	$d[n+1] \otimes \varphi[n]$

were employed, yielding, thus, two audio signal records for the left and right microphones. The sampling frequency f_s was set to 3 kHz. Six reference signals ($J = 6$), i.e., $\mathbf{x}_j[n] = [x_j[n], x_j[n-1], ..., x_j[n-N+1]]^\top$ for $j = 1, 2, ..., J$, were recorded by three accelerometers fixed at crucial positions of the car structure. The input signal $\delta[n]$ is obtained starting from $d[n]$ (both for the left and right microphones). More precisely, two versions of input signal $\delta[n]$ are considered as summarized in Table 9.1, where $\varphi[n]$ is the impulse response of the considered FIR filter.

As one can observe, both for `CASE 1` and `CASE 2`, the input signal $\delta[n]$ can be considered as a filtered (and delayed) version of the input signal $d[n]$. In `CASE 1`, the delay introduced on $\delta[n]$ is of one sample with respect to $d[n]$. When `CASE 2` is considered, the presence of a main echo is taken into account by following FIR filter $\Phi(z)$

$$\Phi(z) = 0.7 + 0.5z^{-3} \ . \tag{9.2}$$

This correspondence between $d[n]$ and $\delta[n]$ can be evaluated also in terms of primary paths $P(z)$ and $\Pi(z)$. In fact, one can observe that for `CASE 1`, the primary path $\Pi(z)$ is $\Pi(z) = P(z) \cdot z^{-1}$ whereas, when `CASE 2` is considered, $\Pi(z) = P(z) \cdot \Phi(z)$. Note that, both cases are examples of causal systems, since $\Pi(z)$ introduces a delay greater than $P(z)$.

An introduction of the anticipating filter z^{+1} can be considered in order to obtain a non-causal system. The non-causal counter part of the setup shown in Table 9.1 can be thus summarized as in Table 9.2.

Table 9.3: Summary table of the considered setups for the fixed observation filter estimation approach with indirect measurements.

Experiment	Scenario	Case	System Causality	Delay n_0 [sample]	$S(z)$	$\Sigma(z)$
1	1	1	YES	0	Sedan	$S(z) \cdot z^{-1}$
	2				Headrest chair	
2	2	2	YES	0	Sedan	$S(z) \cdot z^{-1}$
	2				Headrest chair	
3	1	1	NO	1	Sedan	$S(z) \cdot z^{-1}$
	2				Headrest chair	
4	1	2	NO	1	Sedan	$S(z) \cdot z^{-1}$
	2				Headrest chair	

Two types of secondary paths, i.e., the acoustic channel between the loudspeakers and monitoring microphones, were considered. They are the so-called "sedan" and "Headrest Chair" secondary paths shown in Sections 2.1 and 6.4, respectively. The secondary paths from the loudspeakers to the virtual error microphone, referred to as $\Sigma(z)$, are a one sample delayed version of $S(z)$, i.e., $\Sigma(z) = S(z) \cdot z^{-1}$. The combination of all these setups can be arranged in Table 9.3. For the sake of compactness, only Experiment 1 with Scenario 1 is here considered. Results on Experiment 4 with Scenario 2 are presented in Appendix G.1.1. Similar considerations can be drawn on the remaining experiments.

In Fig. 9.1(a), the impulse responses of observation filter $\mathbf{o}^{\text{opt}}[n]$ and its estimate $\mathbf{o}[n]$, obtained through (8.15) for the `CASE 1` of Table 9.1, are shown. In fact, the orange triangle-marked impulse response, denoting the estimate is quite similar to the blue one. Since the filter to be estimated is a pure delay, the magnitude response is flat (see Fig 9.1(b)). However, an important clarification on the employed setup is here needed. By looking at Fig 9.1(b), one can immediately notice that above 750 Hz the magnitude frequency response $|O(f)|$ of the estimated filter (orange) grows slowly, then falls down. This behavior is due to the employed input signal in this observation filter estimation experiment, that is the input signal $d[n]$ obtained from the acquisition on sedan. The energy contributions of this signal above 750 Hz are very low, and the observation filter can not be consequently well determined. However, since the

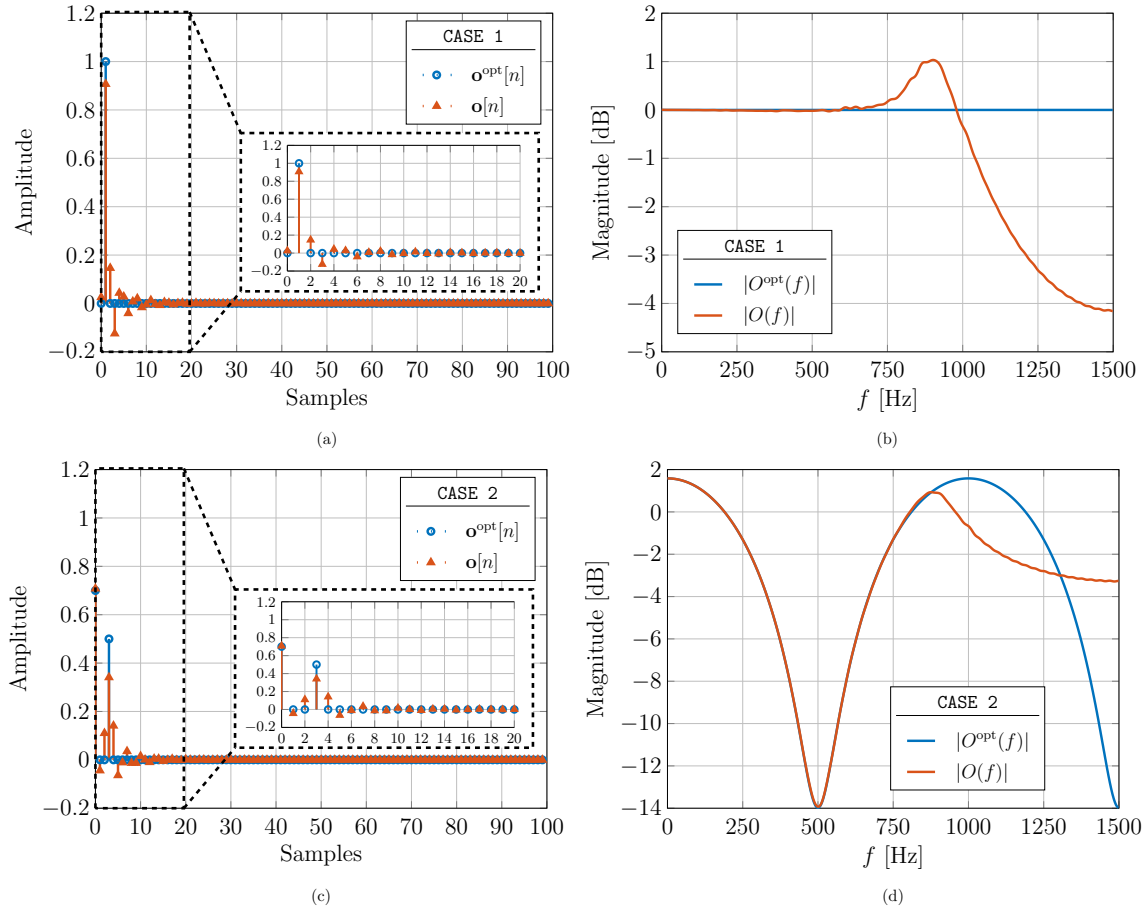


Figure 9.1: Impulse (a) and (c) and magnitude (b) and (d) frequency responses of the optimal observation filter and its estimate for simulation setup `CASE 1` and `CASE 2`, respectively.

secondary paths introduce an anti-aliasing filter with a cut-off frequency at 750 Hz, this inaccuracy on the estimate does not compromise the performance when the ANC system is ON.

Similar considerations can be drawn by observing Fig. 9.1 that represents the impulse (c) and the magnitude responses (d) for the `CASE 2` setup as summarized in Table 9.1. The filter to be estimated is that expressed in (9.2), and, also in this case, it is possible to conclude that the estimation is very well performed.

The SPL spectra as a function of frequency for the single quiet zone, multiple reference SISO ANC when the system is OFF and ON for left and right virtual error microphones are shown in Figs. 9.2(a) and (b), respectively. These results refer to

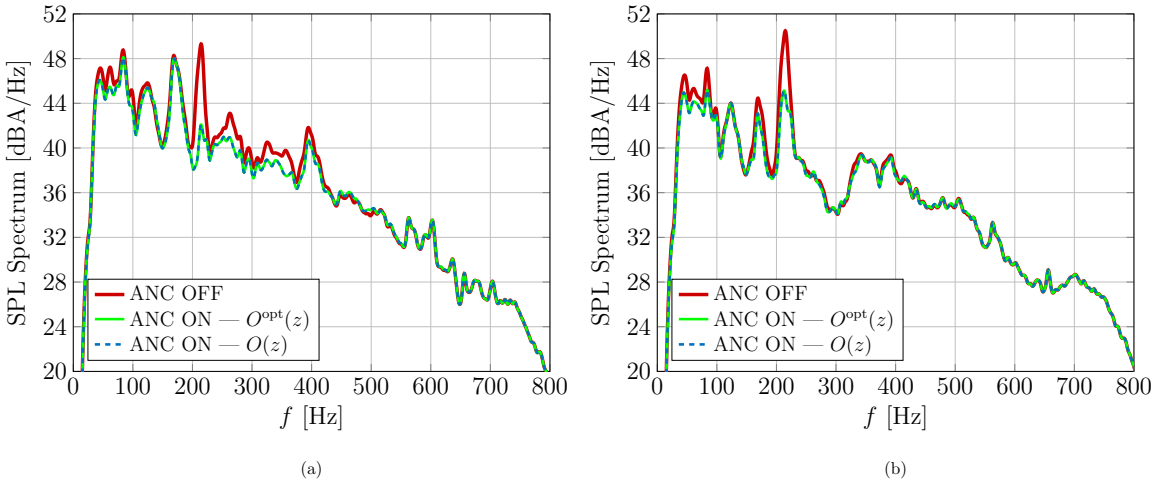


Figure 9.2: SPL spectra for a multi-reference single-quiet-zone SISO system for ANC OFF and ANC ON with optimal and estimated observation filter for Experiment 1 and Scenario 1. Left (a) and right (b) virtual error microphones.

Experiment 1 and Scenario 1 as indicated in Table 9.3. In particular, the SPL spectra for ANC ON is plot both for the case of optimal observation filter $O^{\text{opt}}(z)$ (green curve) and for the estimated observation filter $O(z)$ (blue dotted curve). One can easily notice that, for both virtual error microphones, the maximum peak noise cancellation is experienced around 208–215 Hz, that is the frequency range in which the spectral coherence between disturbance and reference signals is very high, as shown in Part I. This particular frequency range refers to the noise caused by the cavity modes of the tire rolling on the asphalt [5,26]. Moreover it is possible to observe that within 100 Hz a weak noise cancellation is performed for both microphones, whereas only the left one experiences a good cancellation within 250–400 Hz. Finally, by comparing the green curve with the blue dotted one, it is possible to conclude that the observation filter estimation is very accurately performed, since the two curves are overlapped with each other for the whole frequency range.

Sliding window SPL plots against time for left and right virtual error microphones are shown in Figs. 9.3(a) and (b), respectively. For both virtual error microphones, the noise is nicely mitigated and, during the simulation time, the cancellation seems to increase. It is possible to appreciate that the observation filter estimation was well performed.

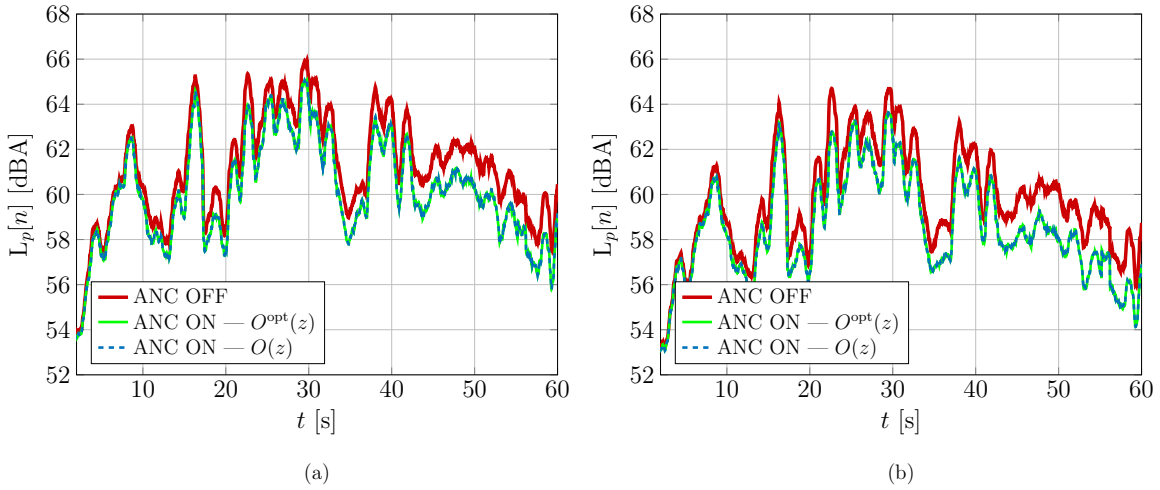


Figure 9.3: Sliding window SPL against time for a multi-reference single-quiet-zone SISO system for ANC OFF and ANC ON with optimal and estimated observation filter for Experiment 1 and Scenario 1. Left (a) and right (b) virtual error microphones.

9.2.1.2 Adaptive Approaches: LMS & RLS

In this section, numerical results obtained with an adaptive observation filter estimation are presented, considering both LMS and RLS algorithms. The employed scheme for these simulations is therefore the one shown in Fig. 8.4.

Differently from the previous analysis, in which observation filter estimation is performed starting from synthesized signals $d[n]$ and $\delta[n]$, here the primary paths $P(z)$ and $\Pi(z)$ are supposed to be available. Thus, desired outputs $d[n]$ and $\delta[n]$ are obtained by filtering white gaussian noise input signal $u[n]$ with the corresponding synthesized primary paths. Sampling frequency was set to $f_s = 48$ kHz. Results can be arranged according to the employed setup as detailed in Tab. 9.4.

For all the considered cases, the employed parameters for LMS and RLS algorithms are reported in Tab. 9.5. The ANC system is supposed to be OFF, and the accuracy of observation filter estimation is assessed only in terms of MSE Υ. An exhaustive and empirical search of the optimum delay n_0 which optimizes system performance is conducted. For the sake of compactness, only the so-called Case 2 is here presented. Further numerical results are shown in Appendix G.1.2.

This setup refers to a system with primary paths $P(z)$ and $\Pi(z)$ that are the same of Case 1 but inverted, as shown in Table 9.4. This yields a non-causal system. The

Table 9.4: Summary table of the considered setups for adaptive observation filter estimation approaches with indirect measurements.

Name	$P(z)$	$\Pi(z)$	System Causality	Filter Length I
Case 0	1	$0.8z^{-1}+0.5z^{-4}+0.3z^{-5}+0.1z^{-8}+0.01z^{-10}$	YES	50
Case 1	$0.8z^{-1}+0.5z^{-4}+0.3z^{-5}+0.1z^{-8}+0.01z^{-10}$	$0.75z^{-2}+0.6z^{-3}+0.4z^{-5}+0.1z^{-7}+0.05z^{-9}$	YES	50
Case 2	$0.75z^{-2}+0.6z^{-3}+0.4z^{-5}+0.1z^{-7}+0.05z^{-9}$	$0.8z^{-1}+0.5z^{-4}+0.3z^{-5}+0.1z^{-8}+0.01z^{-10}$	NO	300

Table 9.5: Employed LMS and RLS algorithm parameters.

LMS		RLS	
Step-size	$\mu = 0.4$	Forgetting factor	$\lambda = 1$
Forgetting factor	$\lambda = 1$	Initial input variance estimate	0.1

observation filter length I was set equal to 300 taps.

MSE performance plots against delay n_0 are depicted in Fig. 9.4(a). For both algorithms, the curves constantly slow down for increasing delay, achieving an interesting flat region within the delay range 75–250 samples. However, outside this range the performance rapidly deteriorates to 0 dB of MSE for delay range within 350–500 samples. The convergence graph of the algorithms, shown in Fig. 9.4(b), demonstrates that the RLS algorithm converges more rapidly than the LMS one, as already observed.

Impulse responses for a delay $n_0^{\text{opt}} = 125$ samples for both proposed algorithms are shown in Fig. 9.5.

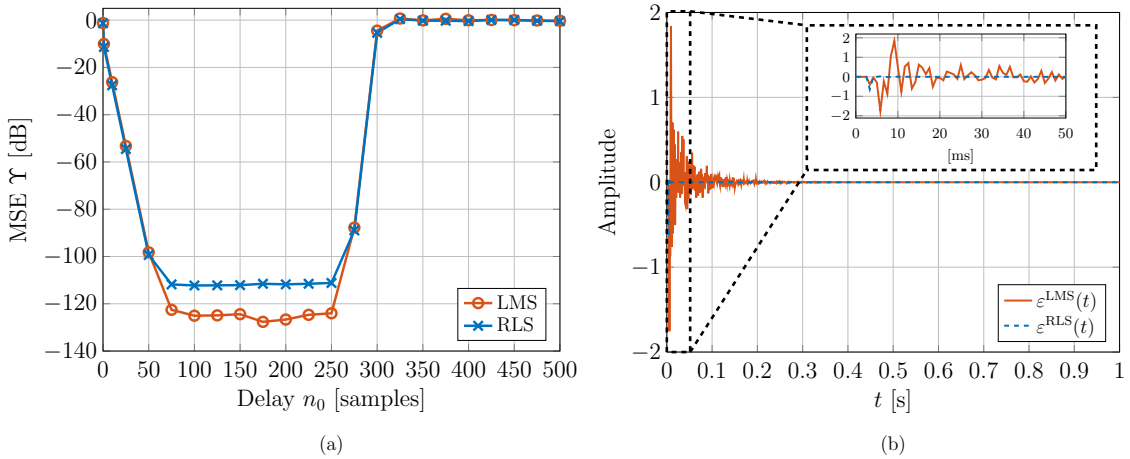


Figure 9.4: MSE against delay n_0 in samples for LMS and RLS algorithms (a). Algorithm convergence speed against simulation time for the case of best delay $n_0^{\mathrm{opt}} = 125$ samples (b).

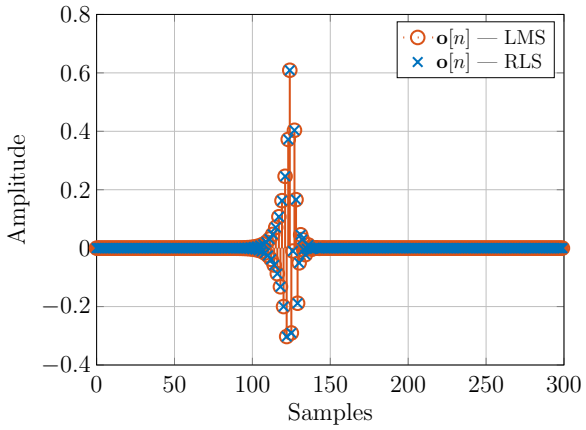


Figure 9.5: Impulse responses of observation filter estimated by employing LMS and RLS algorithms for the best delay $n_0^{\mathrm{opt}} = 125$ samples.

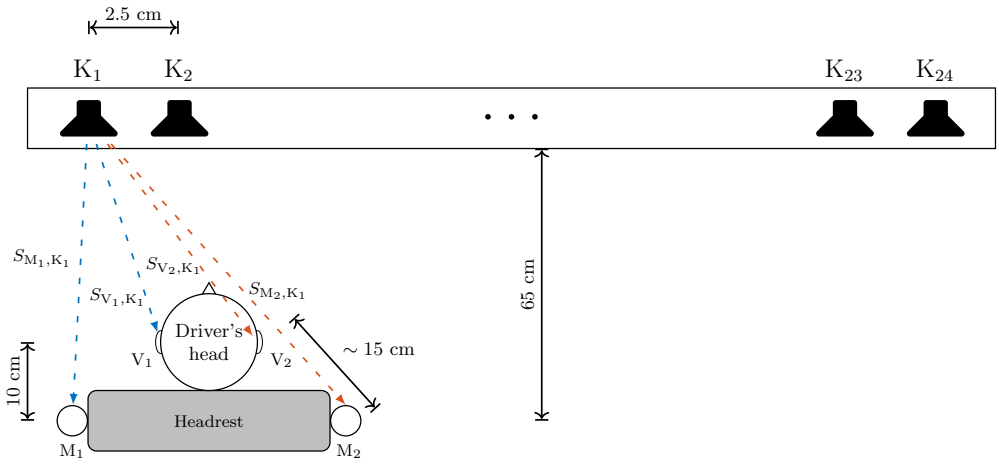


Figure 9.6: Representative scheme of acquired secondary paths on mid-size SUV . Blue and orange arrows denote left and right channels, respectively.

9.2.2 Experimental Response

In this section numerical results on the estimation of the observation filter obtained by employing the experimental primary path filters are discussed. In particular, only the adaptive scheme depicted in Fig. 8.4 was employed for observation filter estimation purposes. White Gaussian Noise (WGN) was used as input signal $u[n]$ with a sampling frequency $f_s = 48$ kHz. Disturbance signals $d[n]$ and $\delta[n]$, thus, are filtered versions of input signal $u[n]$ by transfer functions $P(z)$ and $\Pi(z)$, respectively.

Results can be arranged according to the primary paths $P(z)$ and $\Pi(z)$, obtained from experimental measurements performed by ASK Industries on a mid-size SUV (class J) and in a test garage.

Since this analysis is focused only on the observation filter estimation performance, the ANC system is turned OFF.

9.2.2.1 Mid-size SUV

The results obtained from mid-size SUV measurements are discussed in this section. In Fig. 9.6, the geometrical setup of acquired measurements on the mid-size SUV is depicted. The setup is composed of an array of 24 loudspeakers, spaced by 2.5 cm from each other, placed inside the cabin of the car about 65 cm from the driver's seat. Two binaural microphones, namely V_1 and V_2, are placed at left and right driver's ears,

Table 9.6: Summary table of the considered simulation setups for mid-size SUV experimental filters.

NAME	$P(z)$	$\Pi(z)$	SYSTEM CAUSALITY	FILTER LENGTH I	DELAY n_0
Case 1	$S_{\mathrm{V}_1,\mathrm{K}_1}(z)$	$S_{\mathrm{M}_1,\mathrm{K}_1}(z)$	YES	$100, 200, ..., 600$	$0, 25, ..., 600$
Case 2	$S_{\mathrm{M}_1,\mathrm{K}_1}(z)$	$S_{\mathrm{V}_1,\mathrm{K}_1}(z)$	NO	$100, 200, ..., 600$	$0, 25, ..., 600$

respectively. Finally, two microphones are positioned at the left and right edges of the headrest, i.e., M_1 and M_2, respectively. The distance between binaural and headrest microphones is about 15 cm, that, at a sampling frequency $f_s = 48$ kHz and under the assumption of a speed of sound $v_s = 343.8$ m/s measured at 21°C, corresponds to about 21 samples of delay.

The acquisition method of the acoustic channels between loudspeakers and microphones is the same as that described in Section 6.4, i.e., through the LMS algorithm. These channels are therefore secondary paths, but in this case they are employed as primary paths. The left channels from speaker K_1 (blue arrows in Fig. 9.6) were employed for these simulations, namely, $S_{\mathrm{V}_1,\mathrm{K}_1}(z)$ and $S_{\mathrm{M}_1,\mathrm{K}_1}(z)$. The considered setups are summarized as in Table 9.6. For the sake of simplicity, only Case 2 is here analyzed. The corresponding results for the Case 1 setup are shown in Appendix G.2.1. The MSE measure varies according to the observation filter length I and the delay n_0.

In order to have a simplified overview on performance, a MSE heat color map against filter length and delay is depicted in Figs. 9.7 (a) and (b) for LMS and RLS algorithms, respectively[10]. The larger the MSE value, the warmer the color. For both considered algorithms, the blue area is large, exhibiting remarkable performance improvements. However, in order to obtain a sufficiently good observation filter estimation, at least 600 filter tap weights are needed. By observing Fig. 9.8(a) it is possible to say that, for $I = 600$, the best delay which minimizes the MSE curves (about -11 dB) is $n_0^{\mathrm{opt}} = 325$ samples. Note that the MSE curves decrease according to the increase of delay and, once the minimum is reached, performance gradually degrades. Convergence speed plot, that is shown in Fig. 9.8(b) for both algorithms, exhibits a non-zero residual error, which substantiates the moderate MSE performance. The correspond-

[10]Note that, in the heat map grid, the MSE value is not continuous, but quantized (12 levels).

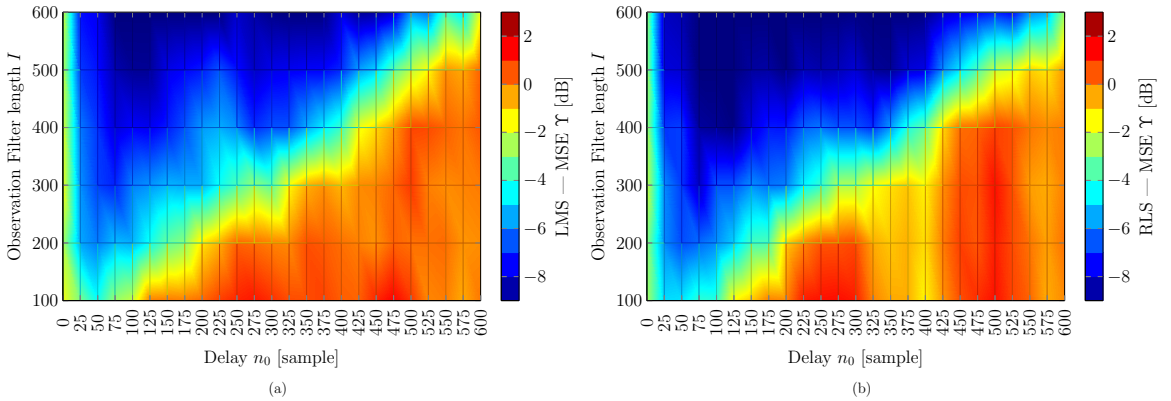


Figure 9.7: MSE heat color map against observation filter length I and delay n_0 in samples for LMS (a) and RLS (b) algorithms for the Case 2 with mid-size SUV experimental responses.

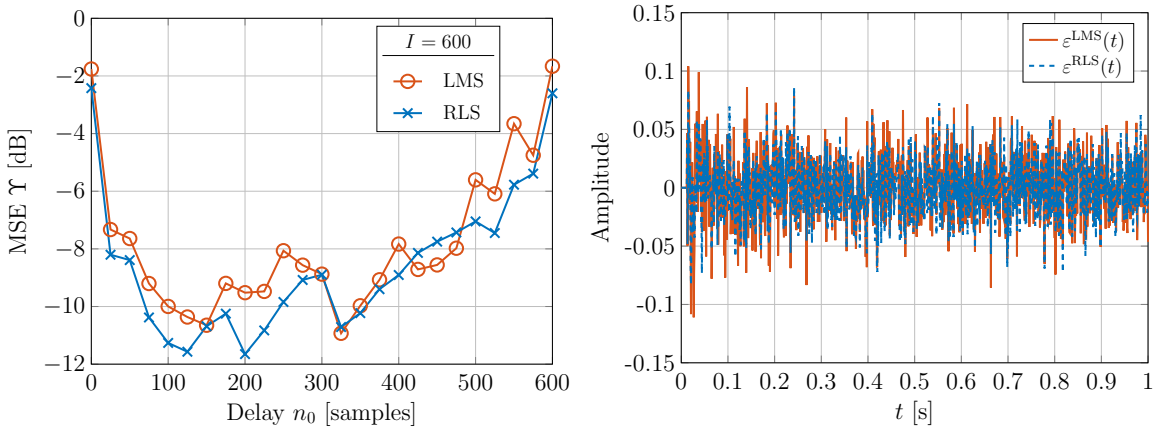
ing impulse responses are shown in Fig. 9.9(a) and (b) for LMS and RLS algorithms, respectively.

9.2.2.2 Garage Testing

Obtained results regarding the experimental setup employing primary paths acquired in the test garage, are presented in this section. The considered setup is shown in Fig. 9.10. Similarly to the mid-size SUV responses, an array of 24 loudspeakers spaced by 4.5 cm to each other, is positioned at a height of 163 cm from the ground. At the same height, but 100 cm far from the loudspeakers, an array composed of 60 microphones is positioned. Each microphone is spaced by 2.5 cm from each other. In order to have a fair comparison with respect to the mid-size SUV case, only the channels between loudspeaker K_{24} and microphones M_1 and M_7, namely $S_{1,24}(z)$ and $S_{7,24}(z)$, are considered. In fact, since the distance between M_1 and M_7 is 15 cm, the same delay of about 21 samples is introduced into the system. The same channel acquisition methodology of the mid-size SUV case is employed for these simulations.

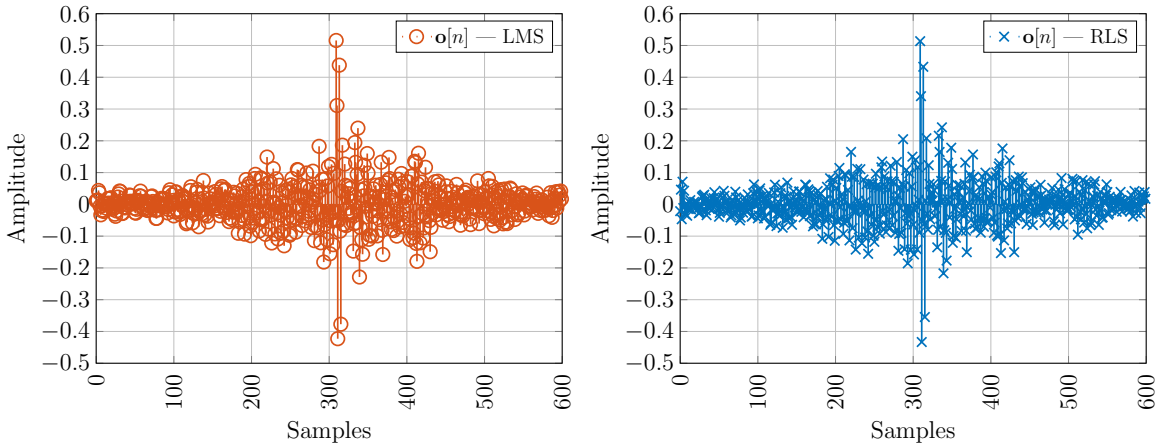
The considered setups are summarized as in Table 9.7. For the sake of compactness, the Case 2, i.e., the non-causal system, is considered. Results on the Case 1 are shown in Appendix G.2.2.

The MSE heat color map against filter length and delay for the $P(z) = S_{1,24}(z)$ and $\Pi(z) = S_{1,24}(z)$ is shown Figs. 9.11(a) and (b) for LMS and RLS algorithms,



(a) (b)

Figure 9.8: MSE against delay n_0 in samples for LMS and RLS algorithms and observation filter length $I = 600$ (a). Algorithm convergence speed against simulation time for $I = 600$ and delay $n_0^{\mathrm{opt}} = 325$ samples (b). Case 2 with mid-size SUV experimental responses.



(a) (b)

Figure 9.9: Impulse responses of the observation filter estimated by employing LMS (a) and RLS (b) algorithms for the $I = 600$ samples and best delay $n_0^{\mathrm{opt}} = 325$ samples. Case 2 with mid-size SUV experimental responses.

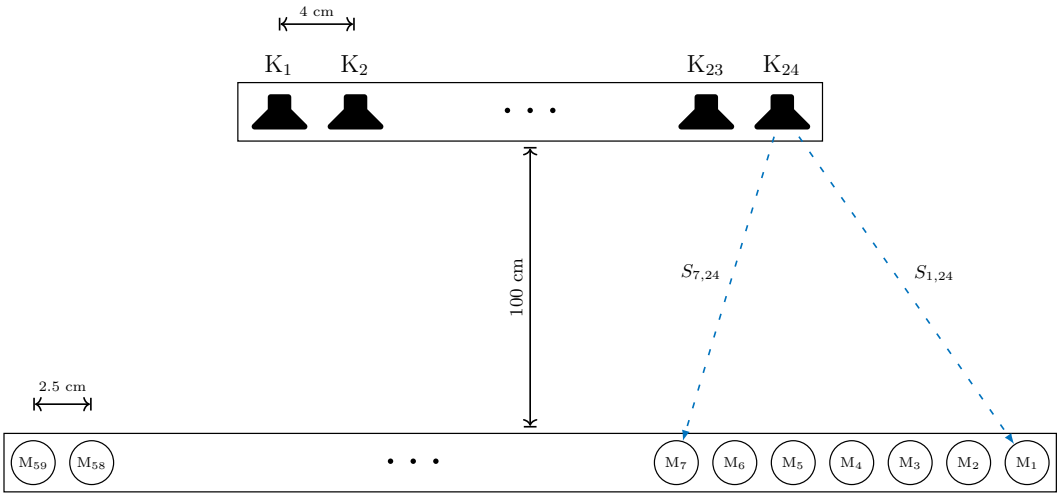


Figure 9.10: Representative scheme of acquired secondary paths in garage testing. Blue arrows denote the considered secondary paths.

Table 9.7: Summary table of considered simulation setups for garage experimental filters.

Name	$P(z)$	$\Pi(z)$	System Causality	Filter Length I	Delay n_0
Case 1	$S_{7,24}(z)$	$S_{1,24}(z)$	YES	$100, 200, ..., 600$	$0, 25, ..., 600$
Case 1	$S_{1,24}(z)$	$S_{7,24}(z)$	NO	$100, 200, ..., 600$	$0, 25, ..., 600$

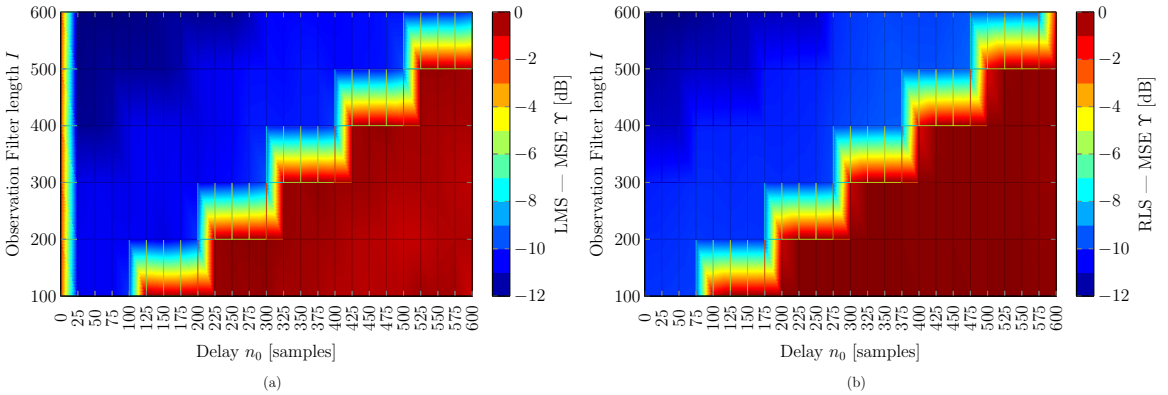


Figure 9.11: MSE heat color map against observation filter length I and delay n_0 in samples for LMS (a) and RLS (b) algorithms for the Case 2 with garage testing experimental responses.

respectively. Note that, once again, a filter length $I = 600$ samples is needed to achieve good system performance.

The MSE behavior against delay in samples, for an observation filter $O(z)$ with a length $I = 600$, is presented in Fig. 9.12(a) for LMS and RLS algorithms. Best performance is obtained for $n_0^{\text{opt}} = 25$ samples, that corresponds to about 0.52 ms of delay regardless of the algorithm strategies. Moreover, note that an almost flat behavior is obtained within the range 300–575 samples of delay. An almost constant residual error for both the considered algorithms can be observed in the convergence speed plot against simulation time in Fig. 9.12(b) for the a filter length $I = 600$ and a delay $n_0 = 25$ samples.

The corresponding impulse responses of estimated observation filter for the best setup are shown in Figs. 9.13(a) and (b) for LMS and RLS algorithms, respectively. Note that both algorithms converge to an almost equal filter, where it is possible to distinguish the direct sound and the reflected one delayed by about 350 samples.

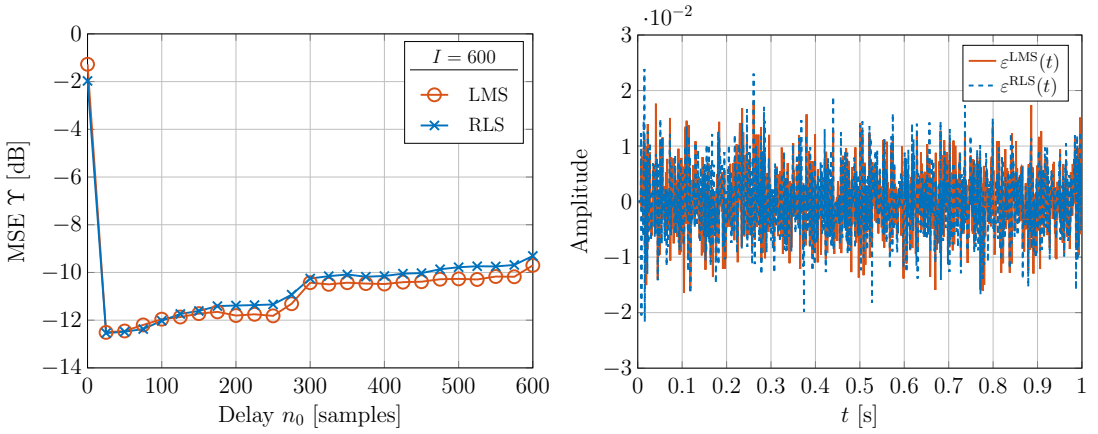


Figure 9.12: MSE against delay n_0 in samples for LMS and RLS algorithms and observation filter length $I = 600$ (a). Algorithm convergence speed against simulation time for $I = 600$ and delay $n_0^{\mathrm{opt}} = 25$ samples (b). Case 2 with garage testing experimental responses.

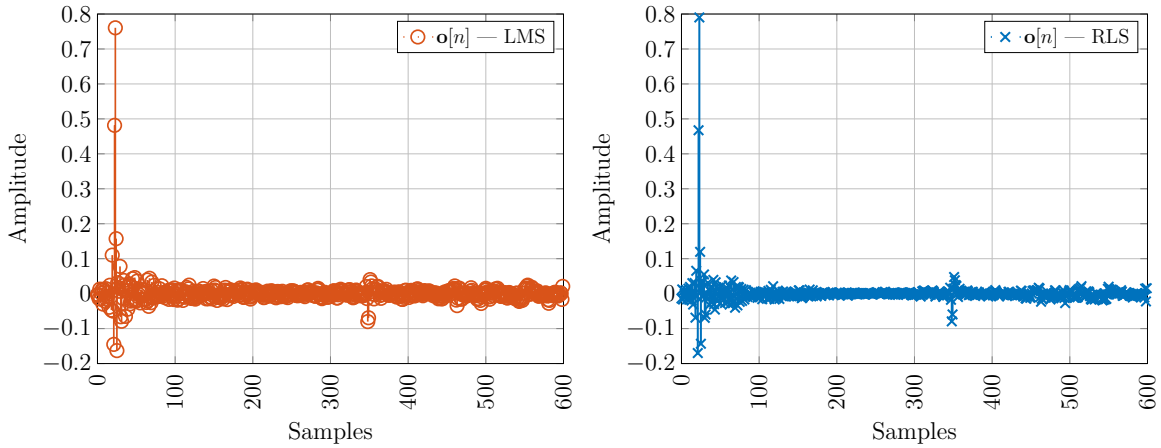


Figure 9.13: Impulse responses of observation filter estimated by employing LMS (a) and RLS (b) algorithms for the $I = 600$ samples and best delay $n_0^{\mathrm{opt}} = 25$ samples. Case 2 with garage testing experimental responses.

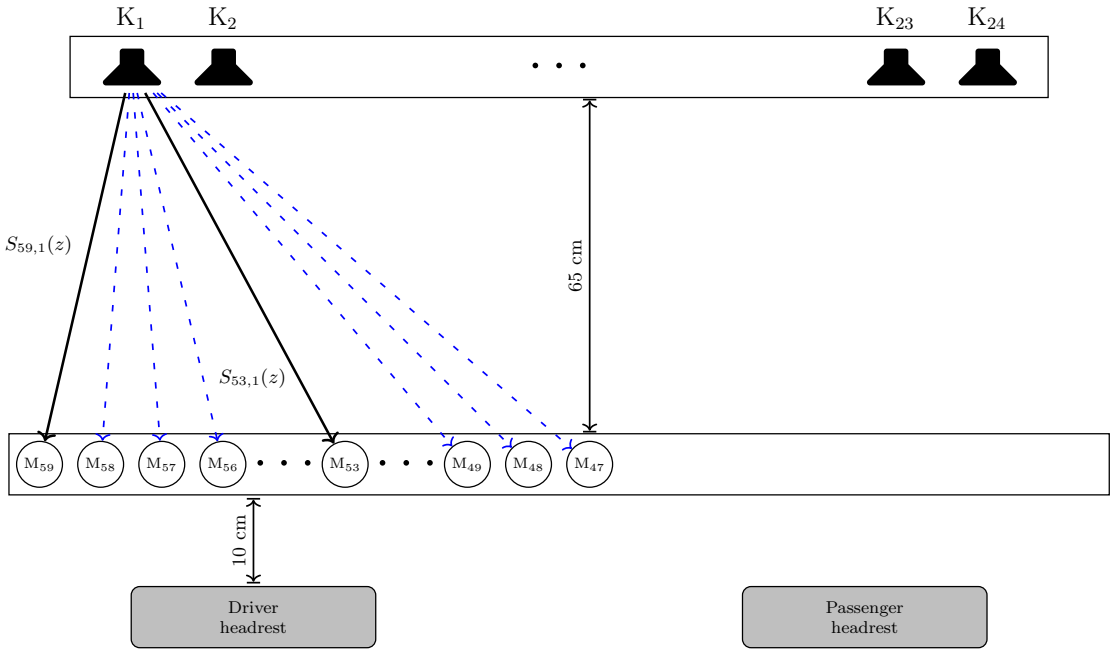


Figure 9.14: Block diagram of mid-size SUV experimental responses setup for OF robustness under virtual microphone mismatch.

9.2.2.3 Microphone Mismatch

Numerical results on observation filter robustness under virtual microphone mismatch are presented in this section. Observation filter estimation is adaptively performed by employing LMS and RLS algorithms. With the aim of saving space, only the mid-size SUV experimental responses are shown. Numerical results for the garage testing experimental responses can be found in Appendix G.2.3.

The considered setup is depicted in Fig. 9.14. Primary paths $P(z)$ and $\Pi(z)$, denoted by black arrows, are represented by the channels $S_{59,1}(z)$ and $S_{53,1}(z)$, respectively. On the contrary, the blue dashed arrows identify the virtual microphone mismatch cases. Hence, it is possible to define the new primary path $\Pi'(z)$ represented by the channels $S_{i,1}(z)$, with $i = 59, 58, ..., 47$ to test the OF estimation robustness under microphone mismatch. Setup parameters can be summarized as in Table 9.8.

The MSE Υ is plotted against microphones for different sub-bands, and shown in Fig. 9.15. In particular, the error signal is narrow-band filtered by 7 approximately equally spaced sub-bands in logarithmic scale (the first and last sub-bands are not to

Table 9.8: Summary table of considered setup parameters for mid-size SUV OF estimation robustness under virtual microphone mismatch.

Description	Value
Observation filter length	$I = 600$
Sampling frequency	$f_s = 48$ kHz
Speaker	K_1
Delay	$n_0 = 25$ samples
Primary path (for monit. mic.)	$P(z) = S_{59,1}(z)$
Monitoring microphone signal	$d[n] = u[n] \otimes p[n] = u[n] \otimes s_{59,1}[n]$
Primary path (for virt. mic.)	$\Pi(z) = S_{53,1}(z)$
Virtual microphone signal	$\delta[n] = u[n] \otimes \pi[n] = u[n] \otimes s_{53,1}[n]$
Primary path (for virt. mic. mismatch)	$\Pi'(z) = S_{i,1}(z)$, with $i = 59, 58, ..., 47$
Mismatched virtual microphone signal	$\delta'[n] = u[n] \otimes \pi'[n] = u[n] \otimes s_{i,1}[n]$

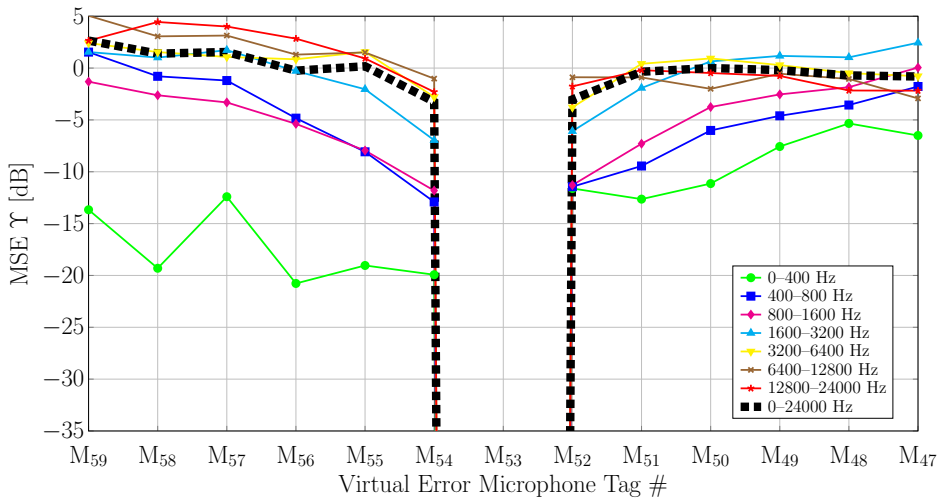


Figure 9.15: MSE Υ per sub-band against virtual error microphones tag number ($\mathrm{M}_{53} \to -\infty$ since $\delta'[n] = \delta[n]$) with mid-size SUV experimental responses.

scale). The dotted-black thick line represents the Υ curve within the overall considered spectrum. In this sense, the dotted-black line can be considered as a weighted average line with respect to the other sub-bands. It is possible to observe that best performance is obtained for the frequency range 0–400 Hz. The green curve, in fact, is always below the curves relative to other sub-bands, achieving the minimum Υ value (about -20 dB) when a mismatch with M_{56} is employed, i.e., $\Pi'(z) = S_{24,56}(z)$ instead of $\Pi'(z) = S_{24,53}(z)$. Note that this microphone mismatch corresponds to a head shift of 7.5 cm. As expected a small head shift corresponds to better performance. When $\Pi'(z) = \Pi(z)$, namely when $\Pi'(z) = S_{24,53}(z)$, Υ tends to $-\infty$ since $\delta'[n] = \delta[n]$. Note that Υ deteriorates with the increase of frequency. This result demonstrates that the observation filter is quite robust to a virtual microphone mismatch for very low frequency ranges; instead it is very weak for high frequency ranges, e.g., above 1600 Hz. Moreover, one can notice that the curves are not symmetric with respect to the distance shift, e.g., for M_{54}, the Υ value is different with respect to the value obtained for M_{52} and so on.

An average value of Υ with respect to the considered microphones for all the sub-bands is calculated as

$$\overline{\Upsilon} = 10\log_{10}\left(\frac{1}{M_{\mathrm{tot}}}\sum_{i=1}^{M_{\mathrm{tot}}} 10^{\frac{\mathrm{MSE_i}}{10}}\right) \qquad [\mathrm{dB}]$$

where, $M_{\mathrm{tot}} = 12$ since the microphone M_{53} was neglected. This analysis gives us information on the sensitivity of the observation filter estimation to the frequency on average with respect to considered microphones. The corresponding bar plot of $\overline{\Upsilon}$ against frequency sub-band is shown in Fig. 9.16.

As previously mentioned, it is possible to observe that best performance is obtained for 0–400 Hz. For increasing frequency, the performance degrades yielding the worst one for 12800–24000 Hz. Surprisingly, one can notice that, the average $\overline{\Upsilon}$ level for the 400–800 Hz window is slightly smaller with respect to 800–1600 Hz. This result may be explained by the fact that some channels may present holes in the spectrum in that frequency range deteriorating performance significantly. For this reason, in Fig. 9.16, a logarithmic trend line is drawn in (dotted-black line). The logarithmic trend line,

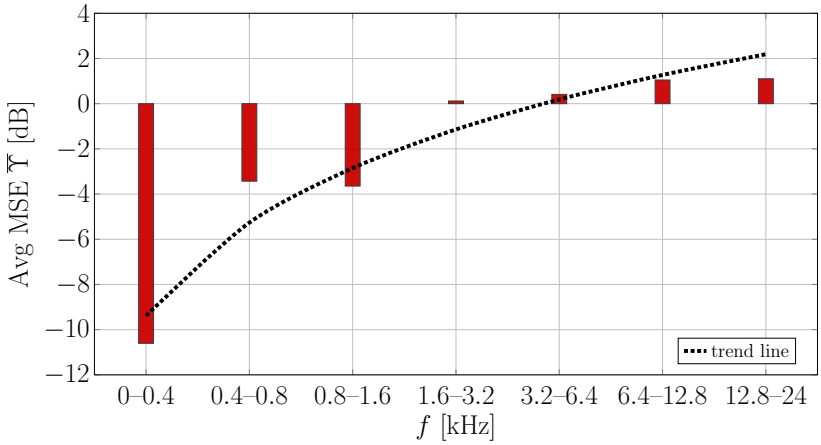


Figure 9.16: Average of Υ with respect to the virtual error microphones against sub-band (M_{53} is not considered) with mid-size SUV experimental responses.

that is calculated as

$$y = 5.9478 \ln(x) - 9.3864$$

confirms what was previously observed.

Finally, a bar plot with the corresponding trend lines of Υ averaged with respect to the symmetric microphones against the considered sub-bands is depicted in Fig. 9.17. As expected, best performance is obtained for low frequency range windowing and few centimeters of microphone mismatch. Green and yellow bars, in fact, exhibit minimum values of $\overline{\Upsilon}$ for all the considered sub-bands. One can note that, only for the sub-bands 400–800 Hz and 800–1600 Hz, the $\overline{\Upsilon}$ values increase with increasing microphone distance mismatch. However, similarly to what was previously said, the corresponding trend lines show the correct curve behavior as frequency increases.

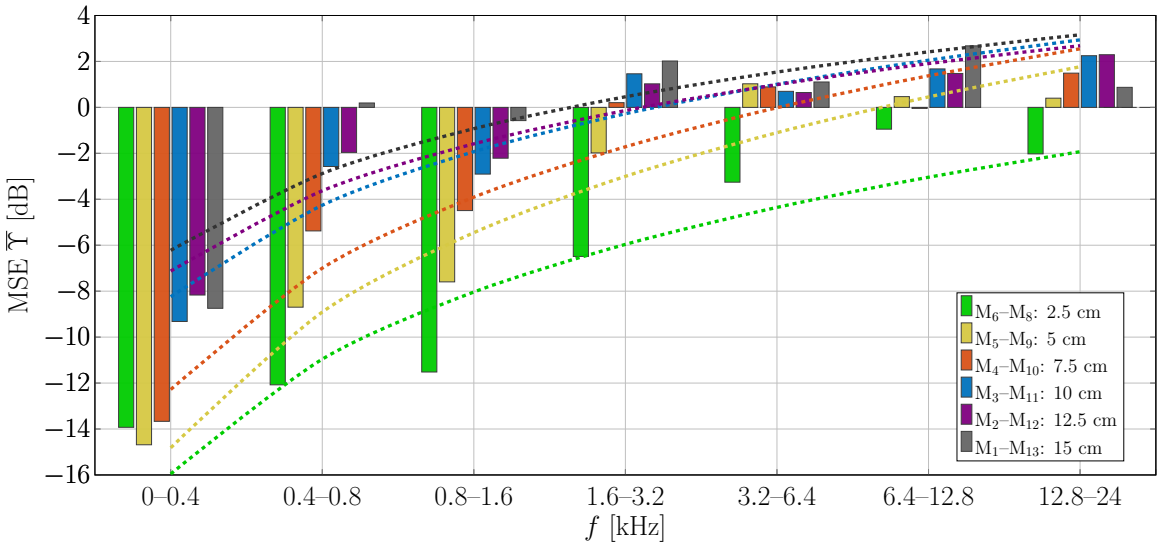


Figure 9.17: Average of Υ with respect to equidistant virtual error microphones against sub-band with mid-size SUV experimental responses. Dashed curves denote corresponding trend lines.

9.3 Direct Measurements

In this section, experimental results on the observation filter estimation accuracy are presented and discussed. As previously mentioned, in this case, the virtual microphone signals are acquired by microphones formerly placed at the desired virtual microphone positions.

During these about 2 years of work in which several experimental setups and parameter configurations were investigated, an extensive campaign of numerical results was obtained, yielding a very large amount of study cases. For this reason, only part of these results is here discussed and presented. For the sake of compactness, in fact, a performance comparison of observation filter estimation by LMS and MMSE algorithms is presented.

Obtained results are organized according to the considered experimental setups, namely, sedan, luxury car and subcompact (B-segment) car in Sections 9.3.1, 9.3.2 and 9.3.3, respectively. More precisely, the most relevant case, consisting of a set of monitoring microphones employed in order to virtualize two positions, is presented in this section for each experimental setup. Note that, for all the considered cases, the ANC system is always OFF, hence, numerical results are assessed only in terms of

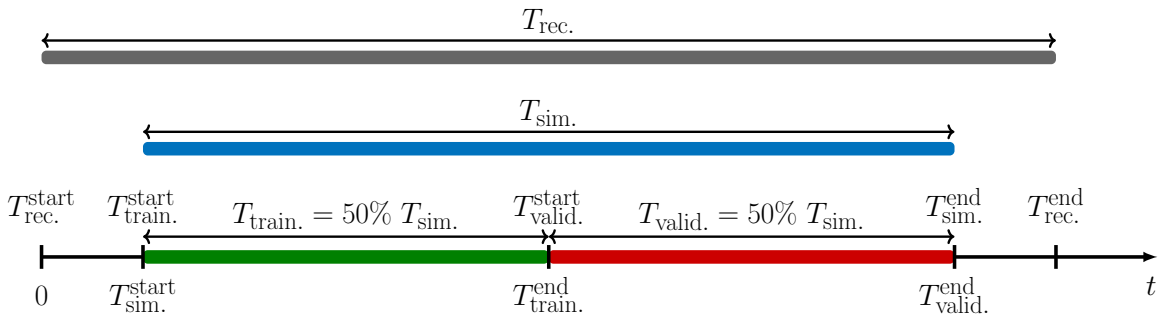


Figure 9.18: Timeline of the simulations.

quality of the observation filter estimation.

Further numerical results on the OF estimation accuracy with direct measurements can be found in Appendix H.

The considered timeline of the simulations is depicted in Fig. 9.18. Microphone signal acquisition starts at $T_{\text{rec.}}^{\text{start}} = 0$ seconds. The estimation of the observation filter is performed according to the adopted algorithm, as described in Chapter 8, within the so-called *training period* $T_{\text{train.}} = T_{\text{train.}}^{\text{end}} - T_{\text{train.}}^{\text{start}}$ seconds. This means that, during the training period, when LMS algorithm is adopted, the algorithm continuously adapts the impulse response of the observation filter based on the time-varying input signal, in order to try to minimize the mean square error. When the fixed algorithm is employed, the training period is used to estimate the vector and matrix according to (8.16) and (8.18), respectively; the observation filter impulse response is computed according to(8.15) and set at the end of this training period. Then, the observation filter is freezed and tested within the so-called *validation period* $T_{\text{valid.}} = T_{\text{sim.}}^{\text{end}} - T_{\text{valid.}}^{\text{start}}$ for the same recorded microphone signals. In particular, for the training and validation periods, the same time duration, equal to the 50% of the overall simulation time $T_{\text{sim.}}$, is spent. Note that, to have more reliable measurements, the first and last portion of the recordings is discarded, yielding thus a simulation time $T_{\text{sim.}} < T_{\text{rec.}}$.

9.3.1 Sedan

The considered signals are obtained from an experimental measurement campaign performed by ASK on a realistic car interior, i.e., that of the sedan. A representative scheme of the monitoring and virtual microphone positions is shown in Fig. 9.19. Two

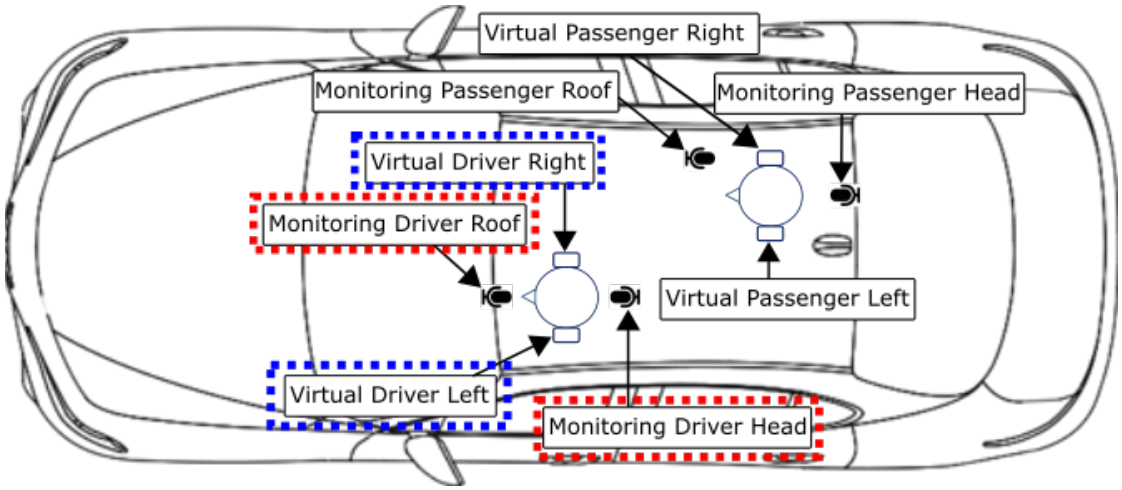


Figure 9.19: Representative scheme of monitoring and virtual microphones within the sedan. Considered monitoring and virtual microphones in red and blue dashed-rectangles, respectively.

monitoring microphones are positioned at the roof and headrest of the car for the driver's seat. Similarly, at the passenger's position, two monitoring microphones are placed at the car internal handle and headrest. Both for the driver and passenger, two binaural microphones located at the left and right ears, are employed in order to acquire the virtual signals $\delta[n]$. The car is run on the road in order to acquire the interior noise of the cabin perceived by the microphones. Note that, for the sake of compactness, only the driver's seat measurements are here considered and analyzed, as depicted by red and blue dashed rectangles in Fig. 9.19.

Disturbing signals were acquired while running the car on two types of road: rough and smooth asphalt. In Figs. 9.20(a) and (b), the picture of the installation of monitoring microphones at the driver and passenger positions are shown, respectively. A picture of the considered rough and smooth roads is shown in Figs. 9.20(c) and (d), respectively. Sampling frequency, was set to 48 kHz. However, with the aim of reducing the computational complexity, a resample operation was performed on the signals yielding $f_s = 12$ kHz. Based on the asphalt type and on the car speed, four different driving scenarios were considered, namely, `Rough 40`, `Rough Variable`, `Smooth 90` and `Smooth Variable` as summarized in Tab. 9.9. For the sake of compactness only scenarios `Rough Variable` and `Smooth 90` are here presented. Corresponding results for the other scenarios are shown in Appendix H.1.

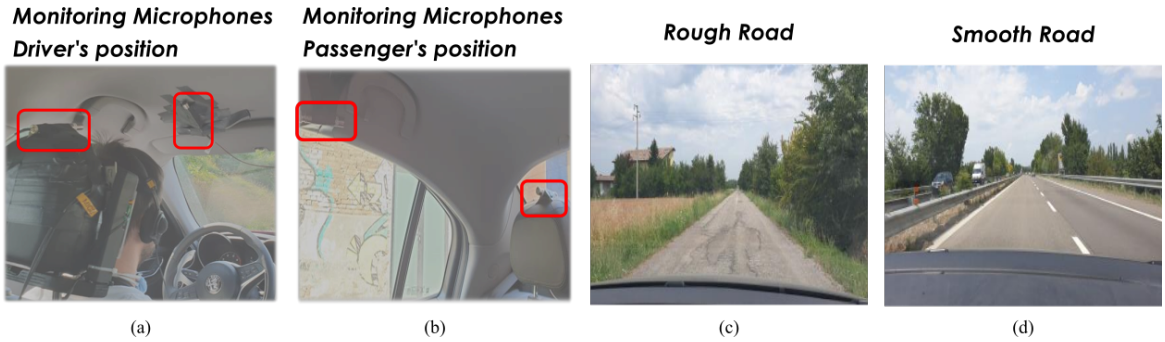


Figure 9.20: Sedan pictures of monitoring and virtual microphone position for driver (a) and passenger (b). Road type pictures: rough (c) and smooth (d).

Table 9.9: Summary table of the road type for sedan direct measurements.

Tag Name	Description
`Rough 40`	Driving on rough asphalt at a constant speed of 40 km/h. Signal acquisition duration: about 80 [s].
`Rough Variable`	Driving on rough asphalt at a variable speed. Signal acquisition duration: about 120 [s].
`Smooth 90`	Driving on smooth asphalt at a constant speed of 90 km/h. Signal acquisition duration: about 120 [s].
`Smooth Variable`	Driving on smooth asphalt at a variable speed. Signal acquisition duration: about 120 [s].

The spectral coherence between the head and roof monitoring microphones and the left virtual one for the `Rough Variable` and `Smooth 90` driving scenario are shown in Fig. 9.21(a) and (b), respectively, in which, for VMT physical limitations, only the intervals from 0 to 1000 Hz are depicted. For both scenarios, it is possible to observe that spectral coherence decreases with increasing frequency and, except for some peak values around 400 Hz and 600 Hz, the main contribution is within 300 Hz. This analysis suggests that the performance of the observation filters is physically limited by the considered microphone positions and may be not sufficiently effective for frequencies higher than 400 Hz.

The impulse responses of the observation filters estimated during the $T_{\text{train.}}$ period by means of the LMS algorithm are depicted in Figs. 9.22(a) and (b) for `Rough Variable` and `Smooth 90`, respectively. The corresponding responses for the MMSE

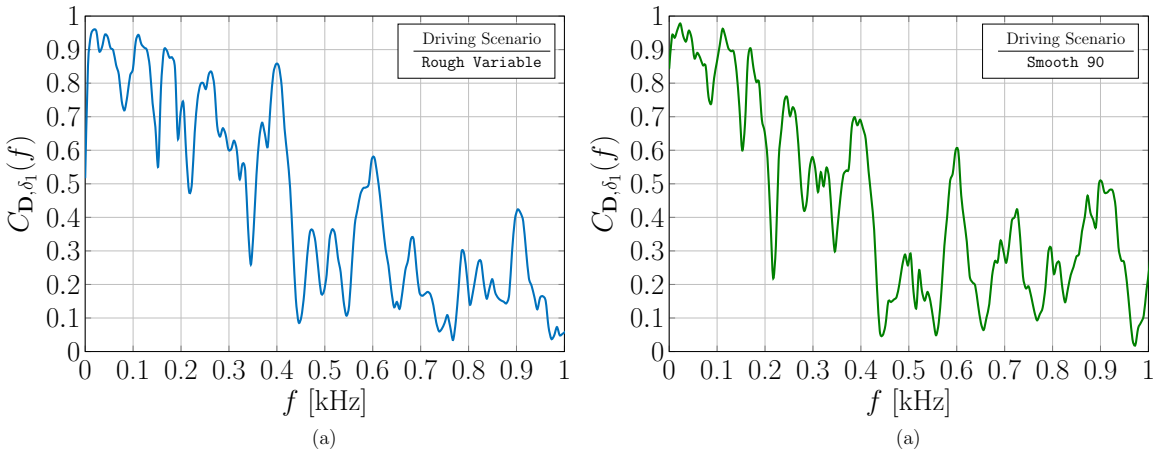


Figure 9.21: Spectral coherence between head and roof monitoring microphone signals and the left virtual one for sedan measurements for `Rough Variable` (a) and `Smooth 90` (b) driving scenarios.

algorithm are shown in Figs. 9.22(c) and (d). In particular, $\mathbf{o}_1[n]$ identifies the estimated observation filter impulse response associated to the acoustic channel from the roof monitoring microphone and the left virtual one, whereas $\mathbf{o}_2[n]$ those from the headrest one. By comparing the two estimation approaches, it is possible to note that both the observation filter length and the delay n_0 are different regardless the driving scenario. In fact, the length which minimizes the MSE of error signal $\varepsilon[n]$, calculated over the validation period, is $I = 700$ and $I = 1500$ samples for LMS and MMSE, respectively. Similarly, the optimal delay is set to $n_0^{\text{opt.}} = 100$ and $n_0^{\text{opt.}} = 200$ as can be observed by the peaks of the impulse responses depicted in Fig. 9.22. By comparing the two approaches, one can notice the difference in terms of amplitude; the impulse response amplitude associated to the MMSE algorithm is almost 10 times larger than the LMS one. This effect may produce a higher SPL of the reconstructed virtual microphone signal $S_{\hat{\delta}}(f)$ for all the frequency range. Moreover, by considering the MMSE approach, some energy contributions are still present at the beginning and at the end of the impulse responses. This may be due to a too short filter length.

The SPL spectra for `Rough Variable` and `Smooth 90` driving scenarios with a driver's left microphone virtualization by means two monitoring microphones placed at the roof and of the driver are shown in Figs. 9.23(a) and (b) for the LMS algorithm and (c) and (d) for the MMSE observation filter estimation, respectively. In particular,

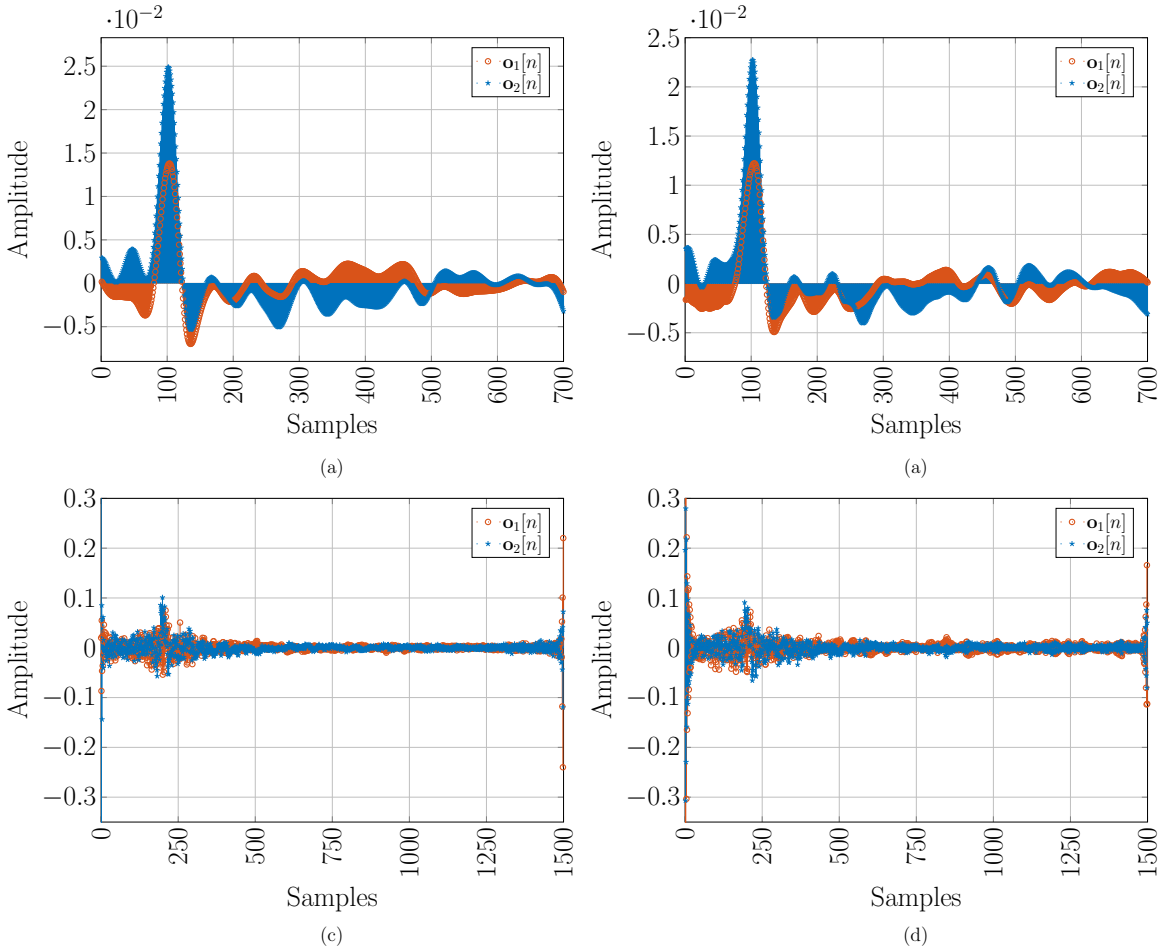


Figure 9.22: Impulse response of estimated observation filter $O(z)$ for driver roof and driver head monitoring microphones and driver's left virtual microphone by means of the LMS algorithm for `Rough Variable` (a) and `Smooth 90` (b) and MMSE algorithms for `Rough Variable` (c) and `Smooth 90` (d) driving scenarios.

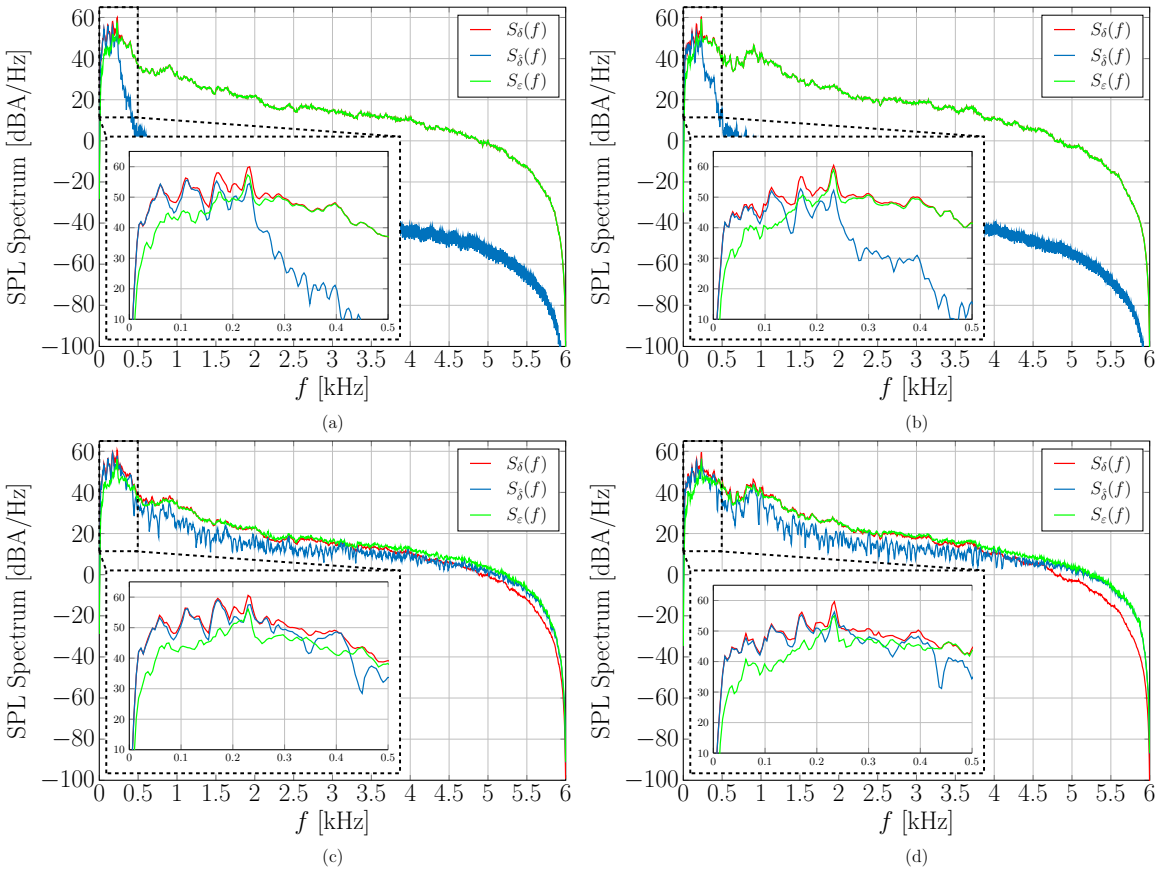


Figure 9.23: SPL spectra of $\delta[n]$, $\hat{\delta}[n]$ and $\varepsilon[n]$ for driver roof and driver head monitoring microphones and driver's left microphone virtualization by means of the LMS algorithm for `Rough Variable` (a) and `Smooth 90` (b) and the MMSE algorithm for `Rough Variable` (c) and `Smooth 90` (d) driving scenarios.

the red line identifies the target signal, i.e., the virtual signal $\delta[n]$, the blue one refers to its reconstructed version $\hat{\delta}[n]$ and the error signal $\varepsilon[n]$ is depicted in green. For all the considered cases, both algorithms show good performance at low frequency regime, i.e., below 200 Hz, since the blue line almost overlaps with the red one. However, it is possible to note that the LMS algorithm is not effective above 250 Hz since the reconstructed signal has basically no energy contribution in this range regardless of the driving scenario. On the other hand, the MMSE exhibits slightly better performance also up to 400 Hz, limit beyond which the observation filter is no longer effective. This analysis suggests that the MMSE algorithm exhibits significant performance gain with respect to its competitor.

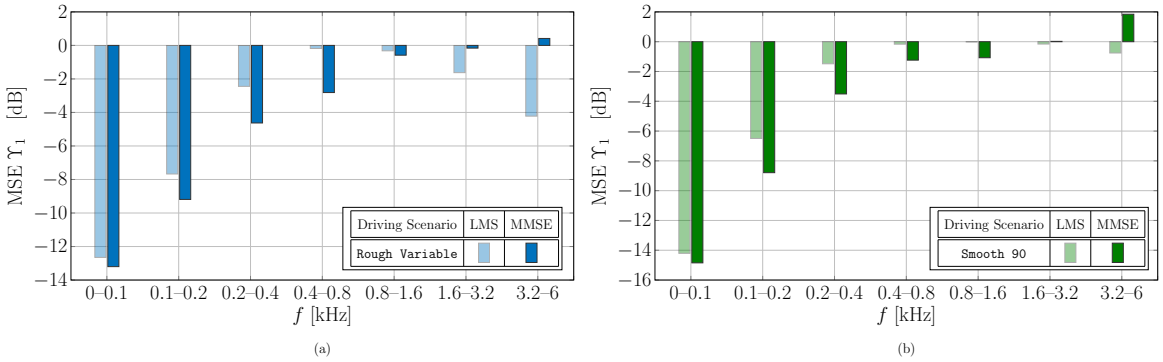


Figure 9.24: MSE against frequency sub-band: comparison between observation filter estimation approaches for `Rough Variable` (a) and `Smooth 90` (b) driving scenarios.

In order to confirm this fact, a MSE analysis of error signal $\varepsilon[n]$ can be performed. The normalized MSE comparison between LMS, depicted in transparent bars, and MMSE, in solid bars, for the reconstruction of the driver's left virtual microphone is shown in Figs. 9.24(a) and (b) for `Rough Variable` and `Smooth 90` driving scenarios, respectively. By analyzing the bar plot it is possible to conclude that both algorithms perform well in the low frequency range, obtaining basically the same MSE values within 0–100 Hz. Performance degrades with increasing frequency, however the MMSE outperforms the LMS algorithm above 200 Hz, indicating that the fixed approach is better. By observing the solid bars in Fig. 9.24, as a consequence of limited computational precision, especially for `Smooth 90` driving scenario, performance degradation is exhibited in the high-frequency regime, i.e., above 1600 Hz. However, since the main energy contribution in the car cabin noise is within low-frequency range (0–500 Hz), this deterioration does not impact the overall system performance. Moreover, one can notice that this degradation is due to the intrinsic characteristic of the estimation approach and it is experienced only in some specific driving scenario (see the Appendix H.1).

Finally, in order to investigate the observation filter estimation accuracy, a robustness test for driving scenario mismatch is performed. More precisely, the observation filters are estimated, by means of the adaptive and fixed approaches, using, in a case, the microphone signals of the `Rough Variable`, and in another, the `Smooth 90` driving scenarios, then these estimates are tested on different input signals deriving from recordings on remaining driving scenarios. Note that, since the time window in which

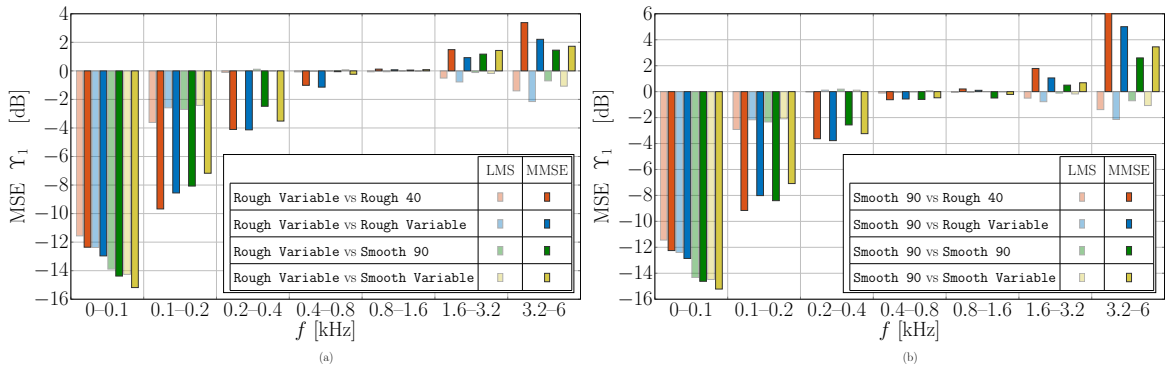


Figure 9.25: MSE against frequency sub-band: comparison between observation filter estimation approaches using the entire duration of `Rough Variable` (a) and `Smooth 90` (b) for driving scenario mismatch.

the MSE evaluation is performed is different, the MSE values for the matched case are slightly different with respect to those shown in Figs. 9.24 for `Rough Variable` and `Smooth 90` scenarios, respectively. The corresponding result in terms of MSE is shown in Figs. 9.25(a) and (b) by comparing the adaptive and the fixed observation filter estimation approaches. In general, it is possible to observe that the MMSE algorithm better performs with respect to the LMS one, regardless of the driving scenario mismatch.

As expected, for both algorithms and driving scenarios, the MSE deteriorates according with increasing frequency, however, it is possible to note that the LMS is not effective above 200 Hz. One cane notice that in several sub-bands the matched case, e.g., `Smooth 90` vs `Smooth 90`, shows a performance worse than the cross ones, regardless the considered estimation approach. This unexpected result may be explained by the fact that signals associated with other driving scenarios may occasionally match the observation filter characteristics and achieve better performance in the specific frequency range. Moreover, even in this case, it is possible to observe that performance deteriorates in the high-frequency regime, showing positive values of MSE. However, this behavior can be ignored since it occurs in a negligible frequency range as previously mentioned.

Table 9.10: Summary table on driving scenarios for luxury car measurements.

Tag Name	Description
`Smooth`	Driving on smooth asphalt at a variable speed 60–90 km/h.
`Rough`	Driving on rough asphalt at a variable speed 60–90 km/h.
`Traffic`	Driving on variable asphalt in traffic condition at a variable speed 30–60 km/h.

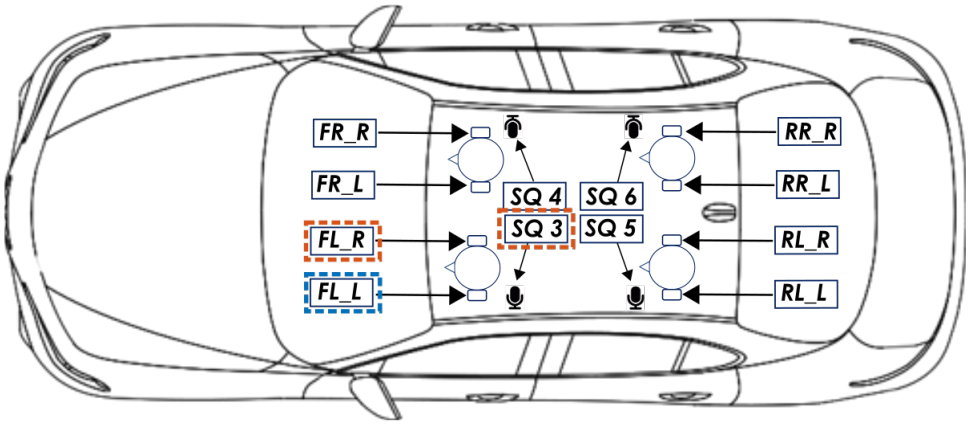


Figure 9.26: Representative scheme of monitoring and virtual microphones within the luxury car. The tags "SQ" identify the monitoring microphones. The blue and orange dashed-line rectangles identify the considered virtual and monitoring microphones positions analyzed in this report, respectively.

9.3.2 Luxury car

Experimental results on the luxury car measurements are shown in this section. The acquisition of microphone signals was performed by ASK Industries on luxury car for about 130 seconds with a sample rate of 48 kHz. With the aim of reducing computational complexity, microphone recordings were down-sampled by a factor $K_{\text{factor}} = 4$, yielding a sample rate of 12 kHz. The driver and three passengers, each one equipped with binaural microphones, allowed to acquire the disturbing audio in the running car in three different driving scenarios: `Smooth` asphalt, `Rough` asphalt and variable asphalt in `Traffic` condition as summarized in Tab. 9.10. A representative scheme of the monitoring and virtual microphone positions is shown in Fig. 9.26. In particular, the tag name of the virtual microphones are labeled in this way: the first letter ("F":

Table 9.11: Summary table on considered scenario for luxury car spectral coherence analysis.

Tag Name	Microphone Tag	# of microphones
$\mathbf{D}_1$	SQ3	1
$\mathbf{D}_2$	SQ3+SQ4	2
$\mathbf{D}_3$	SQ3+SQ4+SQ5	3
$\mathbf{D}_4$	SQ3+SQ4+SQ5+SQ6	4

front - "R": rear) indicates the position of the person within the car, the second letter indicates the side of the person ("L": left - "R": right) and the last one ("L": left - "R": right) identifies the left/right binaural microphone.

Monitoring microphones were labeled as SQ3, SQ4, SQ5 and SQ6 and are installed near the roof grab handle, one for each car door, as depicted in Fig. 9.26. Monitoring microphones were highlighted by orange dashed-line boxes, wheres the virtual one is depicted in blue. One can notice that, due to the lack of monitoring microphones, in this setup, the right binaural virtual microphone of the driver, namely FL_R, was employed as monitoring one.

For the sake of compactness, only results for the `Smooth` road scenario are presented. Further numerical results on different road scenarios can be found in Appendix H.2.

Similarly to what was performed for sedan, a spectral coherence analysis between the monitoring microphones and the driver's left virtual one can be carried out. In particular, in order to investigate the positioning and the number of the available monitoring microphones, namely SQ3, SQ4, SQ5 and SQ6, that are needed to effectively virtualize a position, four different scenarios were considered as shown in Tab. 9.11 (see also Fig. 9.26 as aid). The spectral coherence results for the `Smooth` road scenario are depicted in Fig. 9.27, in which only the first 1000 Hz are shown. Except for some specific peaks, e.g., about 550 Hz and 650 Hz, the coherence degrades with increasing frequency and, as expected, the main coherence contribution is within low-frequency range, i.e., 0–300 Hz. Moreover, it is possible to note that the larger the number of monitoring microphones, the higher the coherence. In fact, best performance is obtained by employing four monitoring microphones, i.e., tag name $\mathbf{D}_4$. However, due to the significant distance of the monitoring microphones from each other and con-

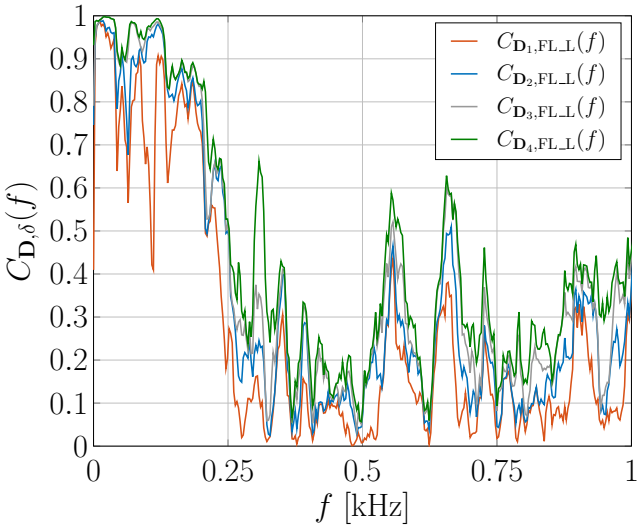
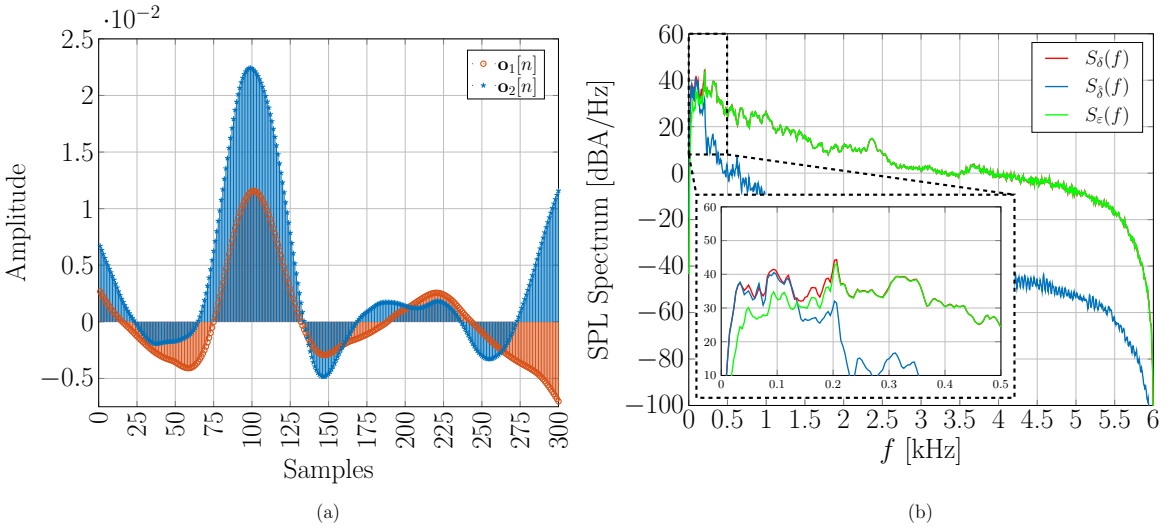


Figure 9.27: Spectral coherence between monitoring microphone signals and the left virtual one for luxury car measurements, for the `Smooth` driving scenario.

sequently between them and the virtual microphone, it is possible to conclude that, unfortunately, even by using all available monitoring microphones, this type of microphone configuration cannot be effective above 300 Hz for an ANC application with VMT. For this reason, the observation filter estimation is performed by employing SQ3 and FL_R as monitoring microphones in order to virtualize FL_L, as previously mentioned. Observation filter estimation was performed employing only the LMS algorithm.

The estimated impulse responses of the acoustic channels from SQ3 and FL_R monitoring microphones to the virtual one (FL_L) by using the LMS algorithm are depicted in Fig. 9.28(a) for the `Smooth` road scenario. In particular, $\mathbf{o}_1[n]$ identifies the observation filter impulse response between the SQ3 and FL_L microphone and $\mathbf{o}_2[n]$ the remaining channel, i.e., between FL_R and FL_L. It is possible to observe that the optimal filter length and the delay which maximize the MSE performance are $I = 300$ and $n_0^{\mathrm{opt.}} = 100$ samples. One can notice that, even in this case, at the end of the impulse responses, energy contributions are present.

In Fig. 9.28(b), the corresponding SPL spectrum is shown. Good estimation performance is obtained at very low-frequency, since the reconstructed version of the virtual signal (in blue) almost overlaps with the target one (in red). This effect is emphasized



(a) (b)

Figure 9.28: Observation filter estimation results for luxury car experimental measurements for `Smooth` driving scenario. Driver's left binaural microphone virtualization by means of monitoring microphone SQ3 and virtual microphone FL_R by employing the LMS algorithm. Impulse responses of observation filters estimated for the best delay $n_0^{\text{opt.}} = 100$ (a). SPL spectra of target signal $\delta[n]$, its reconstructed version $\hat{\delta}[n]$ and their difference $\varepsilon[n]$ (b).

by the error signal depicted in green. As previously suggested by the spectral coherence analysis, unfortunately, microphone virtualization becomes unfeasible above 200 Hz.

System performance in terms of normalized MSE as a function of 1-octave frequency bands is shown in Fig. 9.29 (a). This investigation confirms what was observed in the previous SPL analysis. The virtual microphone reconstruction is effective only in a very low-frequency regime, up to −16 dB of MSE within 0–100 Hz. Performance rapidly decreases with frequency increasing; a deterioration of almost 11 dB is obtained in the next octave band, i.e., 100–200 Hz.

As previously noticed in the algorithms performance comparison for the sedan measurements, we may expect that by employing the MMSE approach, overall system performance improves achieving microphone virtualization in a wider band. However, since the goal of this setup is to try to have a preparatory guideline on the monitoring microphone position and number, a MMSE performance investigation is beyond the scope of this setup.

For the sake of completeness, a robustness observation filters test is performed

for driving scenario mismatch. More precisely, the acoustic channels are estimated by using the `Smooth` driving scenario microphone recordings and, after the training period, the observation filters are employed to virtualize the same position with `Rough` and `Traffic` driving scenario monitoring microphone signals as input. The corresponding normalized MSE performance is presented in Fig. 9.29(b). Note that, even in this case, since the window length of MSE evaluation is different, the MSE values do not coincide with what is observed in Fig. 9.29(a) when the matched case is considered. As expected, the matched case performs better than the cross ones. In general, it is possible to conclude that observation filter estimation is robust up to 200 Hz, since obtained results for the cross case are comparable with the matched ones.

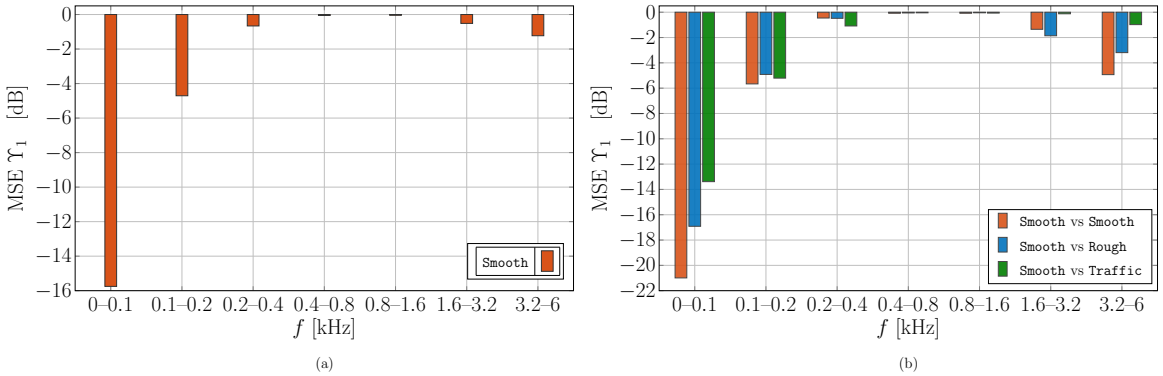
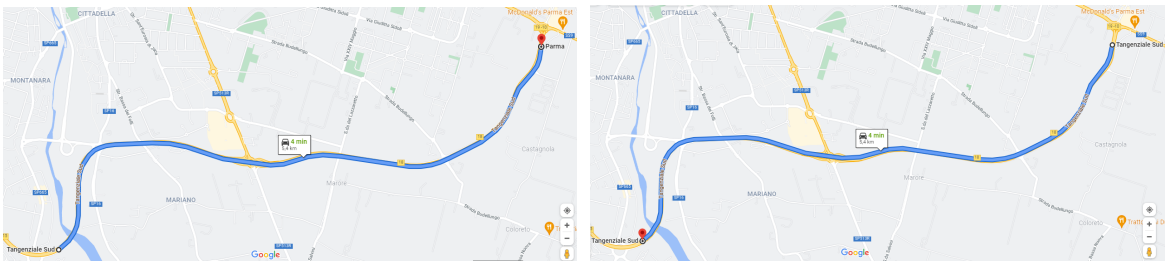


Figure 9.29: MSE against frequency sub-band for luxury car experimental measurements for `Smooth` driving scenario (a) and road mismatch (b).

9.3.3 Subcompact car

The following microphone signals are obtained in an experimental measurement campaign performed on a subcompact car (class B) running on a smooth asphalt of a closed path at a variable speed from 60 km/h to 90 km/h. In particular, the considered closed road trip was made moving away and coming back with respect to our university office in Parma, yielding thus `Forward` and `Return` scenarios, respectively. The trip lasts about 5 minutes with a distance of about 5.4 km per way, as it can be observed by the road map depicted in Fig. 9.30 for `Forward` (a) and `Return` (b) scenarios. The rationale is to estimate the observation filters in one of these two ways and to test them in the remaining one. In fact, although the roads are quite similar with each other,



(a) (b)

Figure 9.30: Road map for subcompact car measurements on smooth asphalt in a ring road at variable speed 60–90 km/h. Moving away (`Forward`) (a) and coming back (`Return`) (b) with respect to the university office in Parma.

weather conditions, during the recordings, changed significantly (it was raining during `Forward` trip and only cloudy during `Return`). This weather event may be very useful in order to effectively check the robustness of the observation filters under unmatched conditions.

Microphone signals were acquired by using a well-known professional portable multi-track field recorder with 8 channels, namely ZOOM F8 produced by ZOOM CORPORATION [56]. The sampling frequency was set to 48 kHz. Since in principle the larger the number of employed monitoring microphones the easier the observation filter task and the better the virtual microphone estimation, the purpose of these measurements is to collect a large number of monitoring microphone signals. In particular, six monitoring microphones were installed within the car cabin. Based on their positioning, two setups were distinguished, namely `Roof` and `Headrest`, as presented in Section H.3.1 and H.3.2, respectively. Moreover, these measurements and the following analysis may be useful to those who are interested in VMT in automotive scenario. In fact, a pragmatic and heuristic approach on positioning and the number of monitoring microphones is suggested in order to effectively virtualize a specific position. These two specific setups (`Roof` and `Headrest`) were considered since they are representative of main applications of interest. In particular, for the `Roof` setup, monitoring microphones were installed at the sun visor and at the roof to evaluate the effectiveness of an array placed above a car occupant's head. Such an array can be effectively placed inside the car structure. Similar considerations may be of interest for the `Headrest` setup, since a microphone array around the occupant's head can be inserted into the seat headrest. Note that, for both setups, the microphone signals are relative to the

driving scenario previously mentioned and depicted in Fig. 9.30.

The purpose of following results is to obtain an effective performance comparison between adaptive and fixed observation filter estimation methods in experimental setups. In particular, implementation guidelines are proposed to obtain a good trade-off between estimation accuracy and material costs. Thus, pragmatic and heuristic approaches on positioning and on the minimum number of monitoring microphones that are needed in order to effectively virtualize a specific position, are suggested.

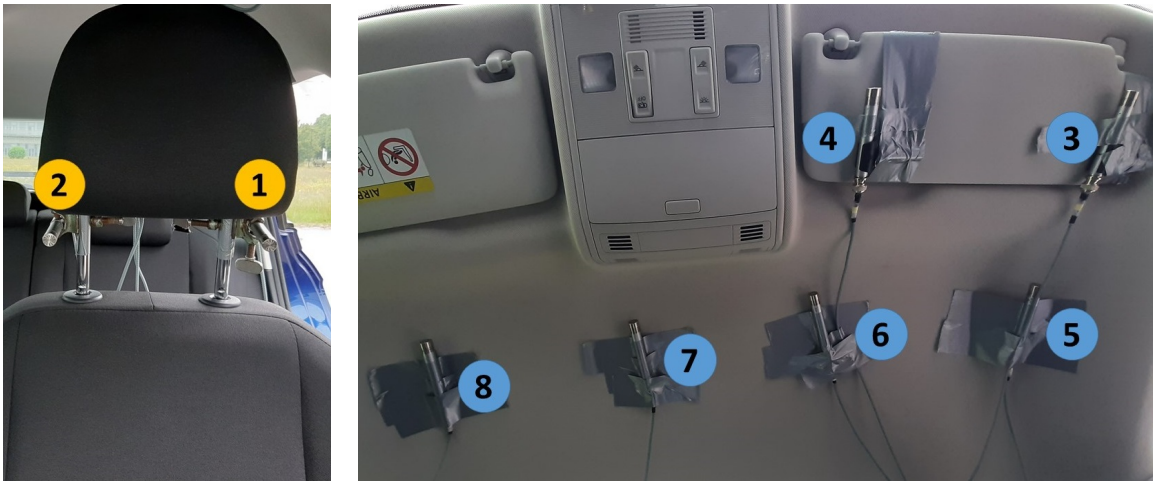
In this case, microphone signals are resampled, from a sampling frequency initially set to 48 kHz, to 6 kHz. For both considered setups, the observation filter estimation is performed by using the `Return` scenario, instead the `Forward` one for test robustness. Moreover, the normalized MSE analysis now is carried on both in 1- and $1/3$-octave bands. Further information on the signal decomposition can be found in Appendix F.

This section is organized as follows. In Section 9.3.3.1, the so-called `Roof` setup is presented and numerical results are discussed. Similarly, the `Headrest` setup is considered in Section 9.3.3.2. Further obtained results by employing these two setups can be found in Appendix H.3.

9.3.3.1 `Roof` Setup

The `Roof` setup consists of a total of six monitoring microphones installed at the cabin roof and at the driver's sun visor. The remaining two channels of the ZOOM F8 were used for the acquisition of the virtual microphone signals around the driver's left and right ears. Fig. 9.31 shows the pictures of the microphone installation within the car cabin. For this setup, free-field Brüel&Kjær microphones for measurements in transport-noise with a sensitivity of 31.6 mV/Pa were employed. In particular, microphones number 1 and 2, depicted in yellow in Fig. 9.31(a), were installed just below the headrest at the maximum possible height. Microphones number 3, 4, 5, 6, 7 and 8, depicted in blue in Fig. 9.31(b), were used as monitoring ones. More precisely, microphones number 3 and 4 were placed at the left and right side of the driver's sun visor, respectively, whereas microphones number 5, 6, 7 and 8 were positioned at the roof of the car, from left to right, respectively.

In Fig. 9.32, a top view representative scheme with the considered distance between microphones for the `Roof` setup is depicted.



(a) (b)

Figure 9.31: Microphone installation within the car interior for the `Roof` setup. Virtual microphones at the driver's headrest (a) and placement of monitoring microphones at the driver's sun visor and roof (b).

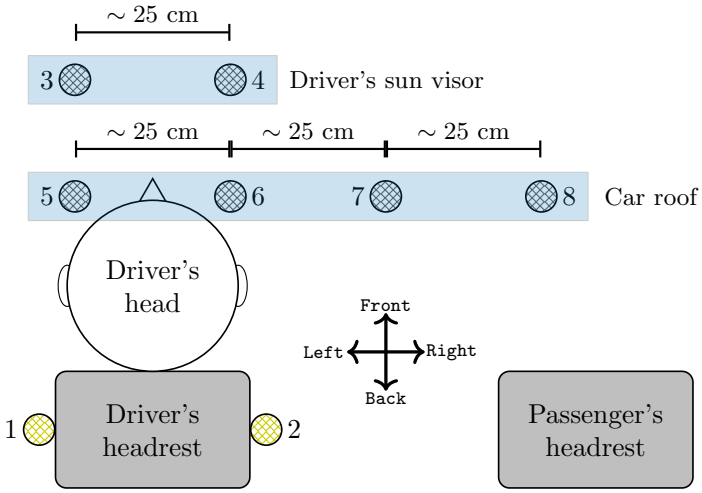


Figure 9.32: Representative scheme (top view) of microphone positioning for the `Roof` setup: in yellow and blue, virtual and monitoring microphones, respectively.

Table 9.12: Summary of the considered scenarios for subcompact car experimental measurements and `Roof` setup.

Scenario	Microphone tag	# of microphones
A	3, 4	2
B	5, 6	
C	3, 4, 5, 6	4
D	5, 6, 7, 8	
E	3, 4, 5, 6, 7, 8	6

In this section, numerical results are discussed for the left virtual position only, i.e., driver's left ear. However, similar conclusions can be drawn also for the right one as shown in Appendix H.3.1.

With the aim of finding the positioning and the number of monitoring microphones that minimize the MSE values, five different scenarios can be identified, as summarized in Tab. 9.12. In particular, in order to virtualize the driver's left ear, scenarios A and B employ only two monitoring microphones placed at the sun visor and at the roof above the driver's head, respectively (see Fig. 9.31 and 9.32). For scenarios C and D, four monitoring microphones, positioned at the sun visor and at the roof near the driver's position and only at the roof are used, respectively. Finally, in scenario E, all six available monitoring microphones are employed.

Figure 9.33 shows spectral coherence of monitoring and left virtual microphone signals for each considered scenario within the frequency range 0–1000 Hz. As expected, coherence degrades with increasing frequency. Moreover, the larger the number of employed monitoring microphones, the better the performance. However, it is possible to observe that, up to 400 Hz, significant coherence levels are exhibited regardless the considered scenario. By comparing the performance for a fixed number of monitoring microphones, one can notice that, placing two transducers at the roof, i.e., scenario B, is potentially more effective than positioning them at the driver's sun visor (scenario A). This may be due to the reduced distance between monitoring microphones and virtual one. When four monitoring microphones are considered, the transducers installation at the roof (scenario C) improves the performance with respect to positioning them at roof and sun visor (scenario D). Finally, as expected, when six microphones are

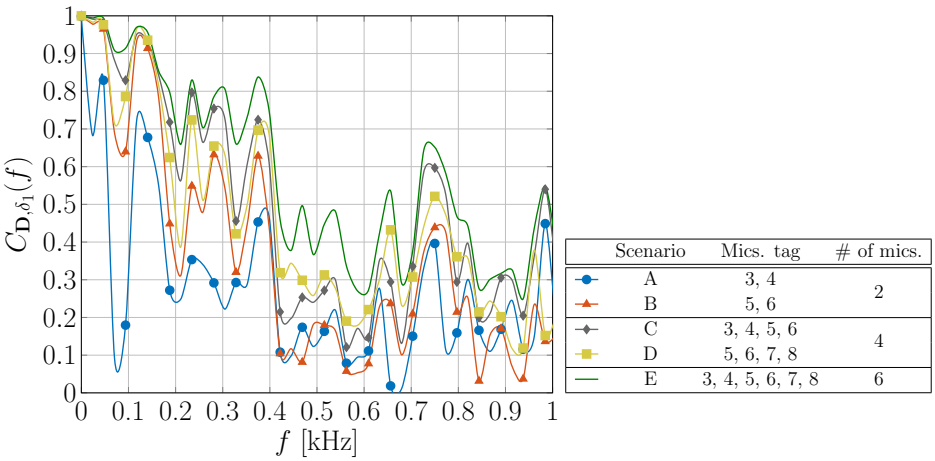


Figure 9.33: Spectral coherence per scenario for the subcompact car experimental measurements and `Roof` setup between monitoring microphone signals and the driver's left virtual one.

available, i.e., scenario E, best performance is obtained.

Performance comparison between adaptive (transparent color bars) and fixed (solid color bars) observation filter estimation, in terms of normalized MSE against 1-octave sub-bands is shown in Fig. 9.34 for each considered scenario, where f_c denotes the center frequency. By fixing the number of employed monitoring microphones, performance shows what was previously suggested, i.e., scenario B performs better than A. Similar MSE values are obtained in scenarios C and D. Although, slight improvement is achieved when monitoring microphones are placed near the driver's position, e.g., scenario C gains almost 1 dB for all the sub-bands, this analysis suggests that placing the transducers at the roof is however a good trade-off. Finally, scenario E performs better than the other ones. By comparing the observation filter estimation approaches, it is possible to conclude that the MMSE algorithm shows significant improvement with respect to the LMS one for all the octave bands regardless the considered scenario. Thus, in general scenarios B, C and E are picked up as the best ones among those with the same number of monitoring microphones.

For these three selected scenarios, a deeper investigation in terms of $^1/_3$-octave bands was performed as depicted in Fig. 9.35. Similar conclusions of the previous analysis can be drawn. In general, it is possible to state that, with respect to the LMS approach, a gain up to 6 dB is achieved when MMSE is employed. Best performance is obtained

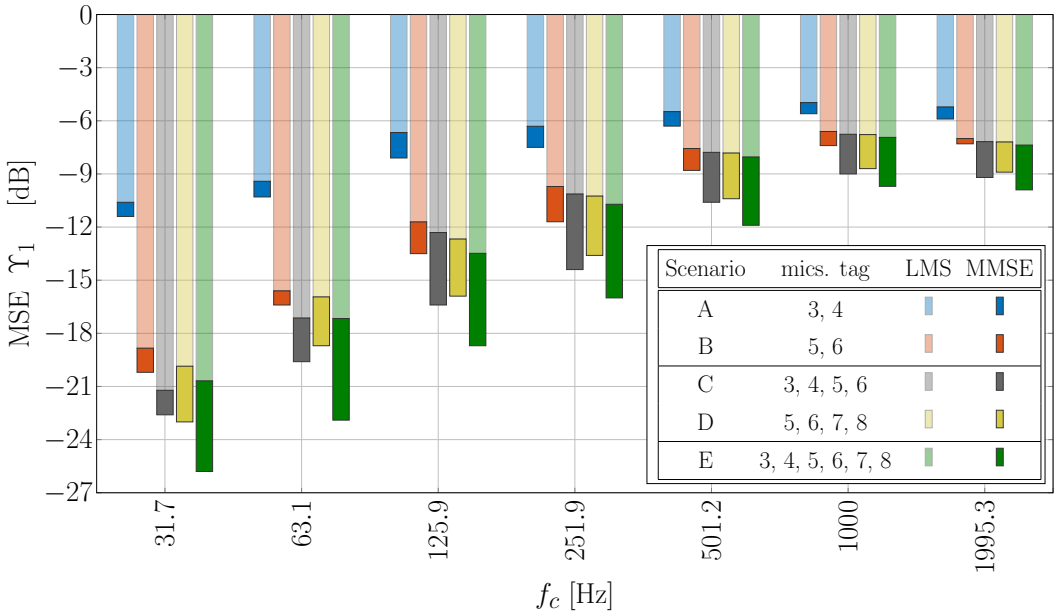


Figure 9.34: Performance comparison in terms of normalized MSE as a function of octave band between adaptive (LMS) and fixed (MMSE) observation filter estimation approaches at the driver's left virtual microphone position for all the considered scenarios with subcompact car experimental measurements and `Roof` setup.

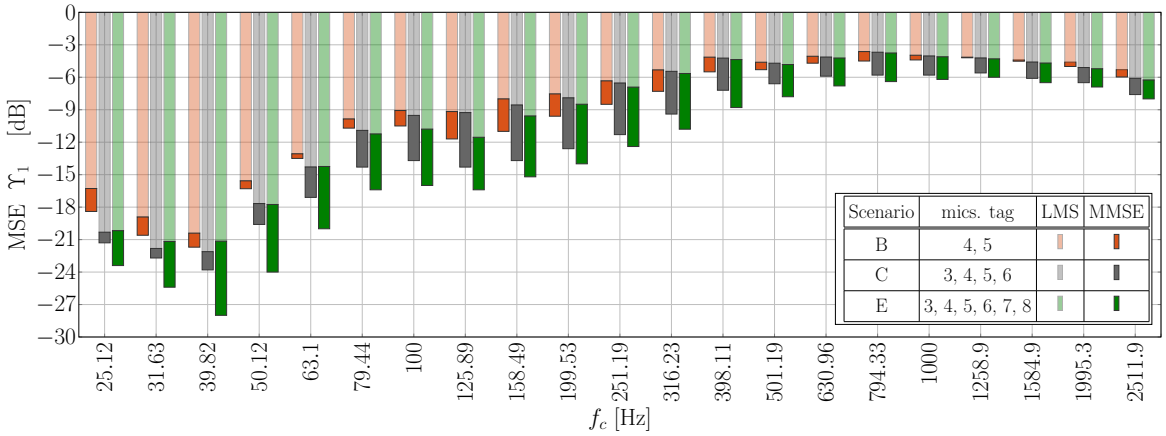


Figure 9.35: Performance comparison in terms of normalized MSE as a function of 1/3-octave bands of adaptive (LMS) and fixed (MMSE) observation filter estimation approaches at the driver's left virtual position for the best scenarios, selected among those with the same number of monitoring microphones, with the subcompact car experimental measurements and `Roof` setup.

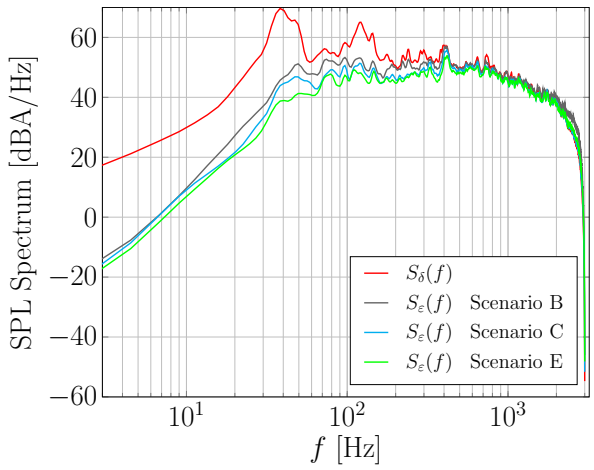


Figure 9.36: SPL spectra of $\delta[n]$ and $\varepsilon[n]$ for the best scenarios, selected among those with the same number of monitoring microphones, for the driver's left microphone virtualization by means of the MMSE algorithm with subcompact car experimental measurements and `Roof` setup.

within the low-frequency range and MSE deteriorates with increasing frequency. However, one can notice that significant improvements with respect to the previous setups presented in Sections 9.3.1 and 9.3.2, are achieved since observation filter estimation is effective in a wider band.

The corresponding SPL spectrum in shown in Fig. 9.36 for the MMSE observation filter estimation. In particular, the SPL spectrum of the target signal $\delta[n]$, in red, and of the error signal $\varepsilon[n]$ for scenarios B (gray), C (cyan) and E (green) is shown. Note that, the x-axis is in log-scale. As previously observed, since the lower the SPL level of the error signal, the better the performance, best behavior is achieved within the low-frequency range.

Finally, the robustness test of the observation filter estimation accuracy for road mismatch is performed for both virtual microphones, i.e., driver's left and right ear. The results are depicted in Fig. 9.37 in terms of normalized MSE against 1/3-octave band for the MMSE approach for scenarios B, C and E. The estimation accuracy deteriorates with increasing frequency. Conversely, performance improves when a large number of monitoring microphones is used. One can notice that the right virtual microphone shows slightly better performance with respect to the left one. This behavior may be explained by the fact that, since the main noise components in the car interior is given

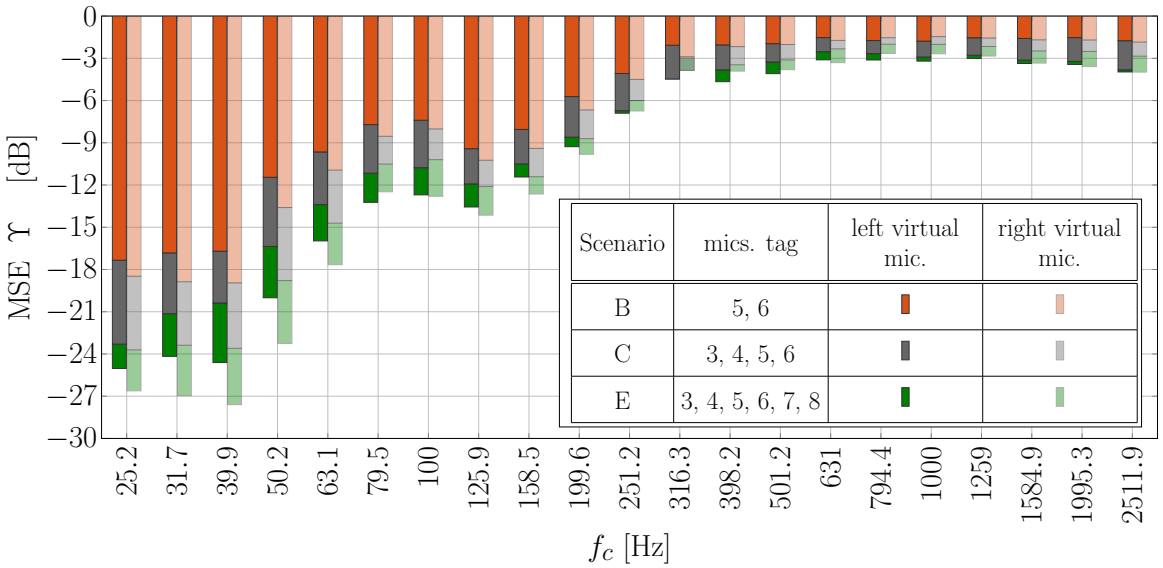


Figure 9.37: Performance comparison in terms of normalized MSE measured over the entire duration of recorded signals as a function of 1/3-octave bands with MMSE observation filters estimation at the driver's left and right virtual positions for the best scenarios selected among those with the same number of monitoring microphones, for road mismatch with subcompact car experimental measurement and `Roof` setup.

by window noise (left side of the driver), the left virtual microphone signal can be more noisy, complicating thus, the estimation task.

Table 9.13 summarizes the considered scenarios (see also Figs. 9.31 and 9.32).

9.3.3.2 `Headrest` Setup

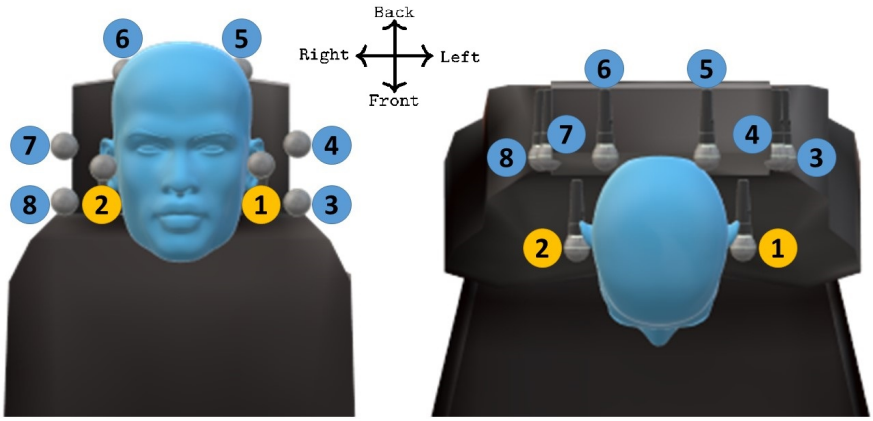
Pictures of the microphone installation for the `Headrest` setup are shown in Figs. 9.38. A Sennheiser manikin positioned at the front passenger's seat was equipped with two binaural microphones, with a sensitivity of 10 mV/Pa, inserted within the manikin's ears in order to collect the perceived virtual microphone signals, as shown in yellow in Fig. 9.38(a). Monitoring microphones were symmetrically positioned around the perimeter of the headrest as can be seen, in blue, in Fig. 9.38(b). In particular, from passenger's left to right sides, microphones number 3 and 8 were placed at the base of the headrest, microphones number 4 and 7 at half height of the headrest, as depicted in the Figs. 9.38. Finally, microphones number 5 and 6 were installed above the headrest, respectively. Note that the employed monitoring microphones are the same as those

Table 9.13: Summary on the considered scenarios in experimental `Roof` setup. The employed parameters for LMS and MMSE algorithms are also shown.

Scenario	M	Microphone Tag	LMS		MMSE	
			n_0	I	n_0	I
A	2	3,4	175	750	200	750
B		5,6	100	500	75	500
C	4	3,4,5,6	100	500	200	1500
D		5,6,7,8	100	500	200	1500
E	6	3, 4, 5, 6, 7, 8	200	500	200	1500

(a) (b)

Figure 9.38: Microphone installation within the car interior for the `Headrest` setup. Virtual microphones at the passenger's ears and monitoring microphones at the sides of the headrest (a). Monitoring microphones positioned around the passenger's headrest perimeter (b).



(a) (b)

Figure 9.39: Representative scheme of microphone positioning for the `Headrest` setup. In yellow and blue virtual and monitoring microphones, respectively. Front view (a), top view (b).

used in the `Roof` setup. A simplified representation of the considered `Headrest` and the corresponding microphone tags is depicted in Fig. 9.39(a) front and (b) top view.

The following results only refer to the passenger's right ear, since it represents the position of main interest for the reasons previously explained. For the sake of completeness, Appendix H.3.2 is dedicated to the results on the left virtual microphone.

Six scenarios are considered in order to optimize the performance in terms of number and microphone positioning on the passenger's headrest. Considered scenarios are

Table 9.14: Summary of the considered scenarios for subcompact car experimental measurements and `Headrest` setup.

Scenario	Microphone tag	# of microphones
A	3, 8	2
B	4, 7	
C	5, 6	
D	3, 4, 7, 8	4
E	4, 5, 6, 7	
F	3, 4, 5, 6, 7, 8	6

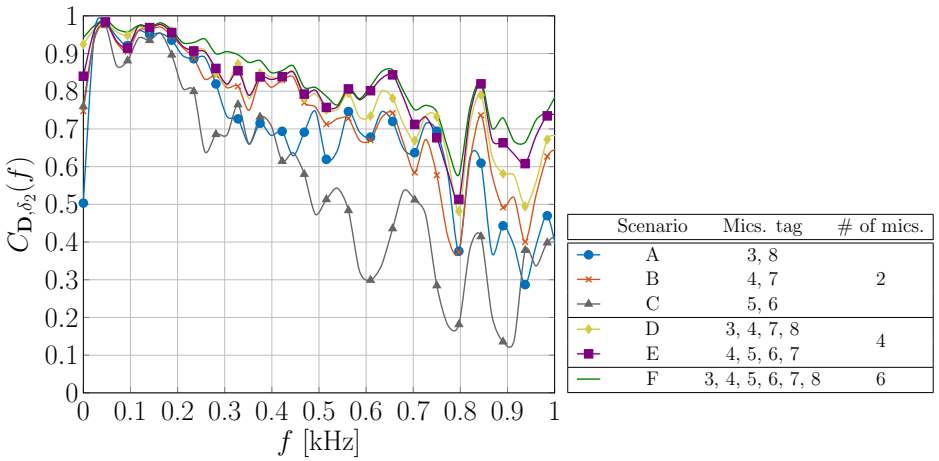


Figure 9.40: Spectral coherence per scenario for the subcompact car experimental measurements and `Headrest` setup between monitoring microphone signals and the passenger's right virtual one.

summarized in Tab. 9.14 (see Figs. 9.38 and 9.39 as aid). Per each scenario, spectral coherence for the right virtual microphone is depicted in Fig. 9.40 within the first 1000 Hz. It is possible to easily observe that better performance with respect to the `Roof` setup is obtained thanks to the reduced distance between monitoring and virtual microphones. In fact, appreciable coherence level is exhibited up to 600 Hz and 1000 Hz when four and six monitoring microphones are employed, respectively. By comparing scenarios with fixed number of monitoring microphones, it is possible to conclude that when two transducers are used, best performance is achieved by positioning them at half-height of the headrest (scenario B). If four microphones are available, a good installation setup is represented by scenario E, i.e., positioning them at half-height and above the headrest. Finally, as expected, best performance is obtained when six monitoring microphones are employed (scenario F).

Results on the normalized MSE against 1-octave bands for LMS and MMSE observation filter estimation algorithms are depicted in Fig. 9.41. Significant wide-band performance is exhibited by the MMSE algorithm with respect to the LMS one. In fact, MSE dB values are doubled, e.g., at 1000 Hz when scenario F is considered. This means that the MMSE is the best observation filter estimation approach. In general, it is possible to identify the best scenarios as B, E and F for two, four and six available monitoring microphones.

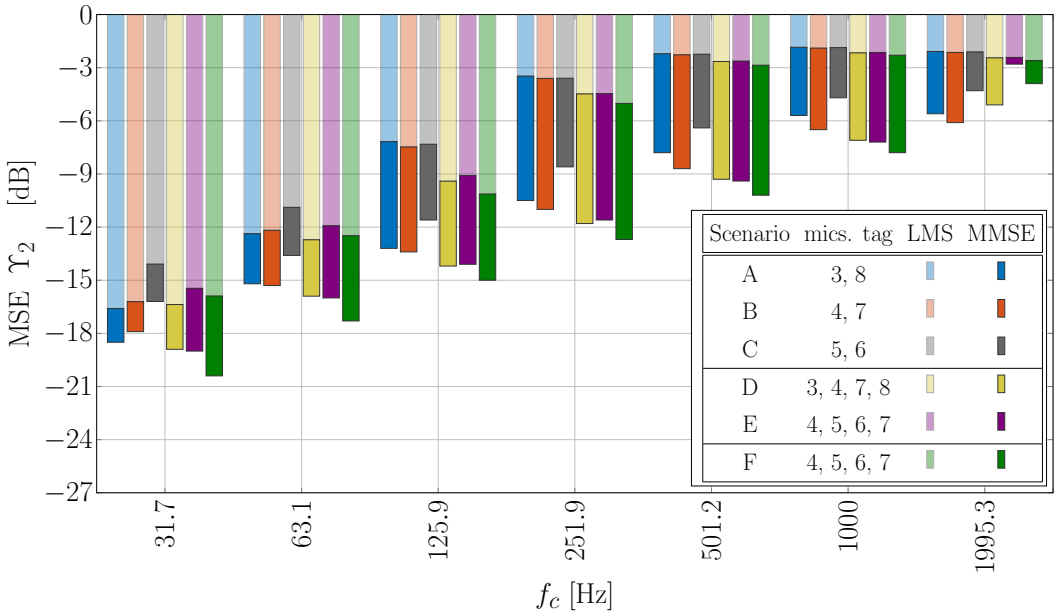


Figure 9.41: Performance comparison in terms of normalized MSE as a function of octave band of adaptive (LMS) and fixed (MMSE) observation filter estimation approaches at the passenger's right virtual microphone position for all the considered scenarios with subcompact car experimental measurements and `Headrest` setup.

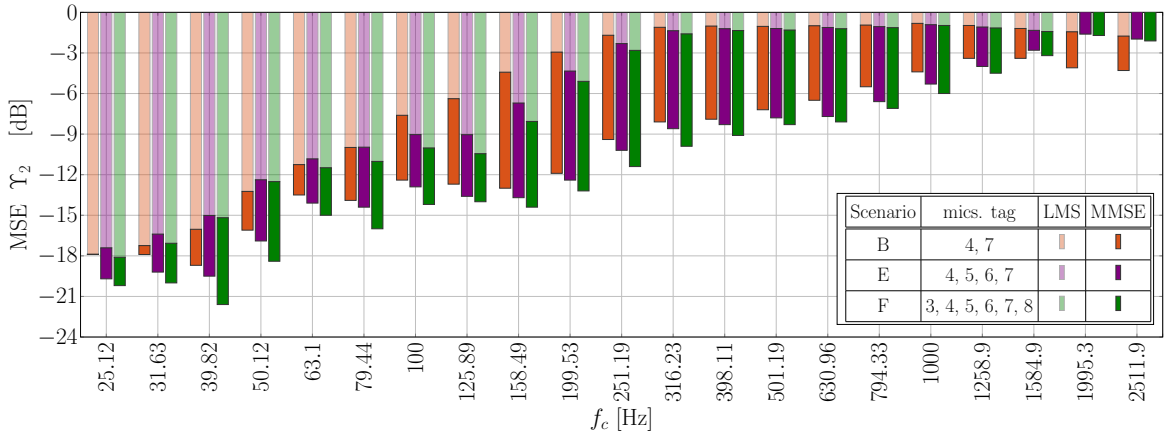


Figure 9.42: Performance comparison in terms of normalized MSE as a function of 1/3-octave bands of adaptive (LMS) and fixed (MMSE) observation filter estimation approaches at the passenger's right virtual position for the best scenarios, selected among those with the same number of monitoring microphones, in the `Headrest` setup.

For these scenarios, further numerical results in terms of normalized MSE against 1/3-octave bands can be found in Fig. 9.42. The MMSE approach shows significant improvements with respect to the fixed one, e.g., almost three times better (in dB) up to 400 Hz, regardless the scenario. Wider band improvements are also obtained with respect to the `Roof` setup. Similar conclusions can be drawn by observing Fig. 9.43. In fact, the SPL spectrum of the error signal is always lower than the target up to 1000 Hz.

Finally, by comparing `Roof` and `Headrest` setups, it is possible to observe that the observation filter performs well in a wider-band when the monitoring microphones are installed at the headrest, since it is effective almost up to 1000 Hz. It is thus clear that an ANC system with VMT can be efficiently realized by implementing a headrest setup similar to this [57].

As in Table 9.13, optimal values of n_0 and I are listed in Table 9.15.

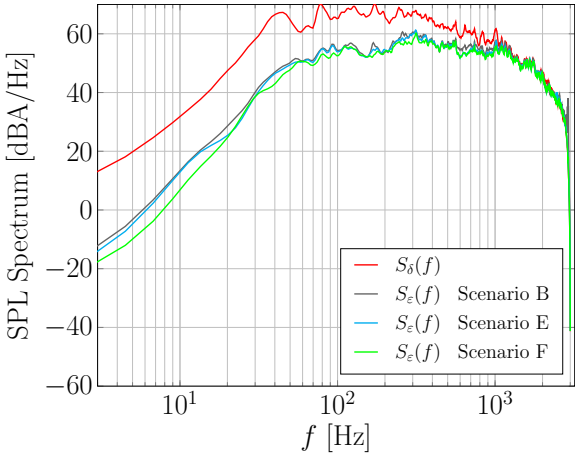


Figure 9.43: SPL spectra of $\delta[n]$ and $\varepsilon[n]$ for the best scenarios, selected among those with the same number of monitoring microphones, for the passenger's right microphone virtualization by means of the MMSE algorithm with subcompact car experimental measurements and `Headrest` setup.

Table 9.15: Summary on the considered scenarios in experimental `Headrest` setup. The employed parameters for LMS and MMSE algorithms are also shown.

Scenario	M	Microphone Tag	LMS		MMSE	
			n_0	I	n_0	I
A	2	3,8	150	500	75	1000
B		4,7	150	500	75	750
C		5,6	125	500	200	1000
D	4	3,4,7,8	150	500	75	750
E		4,5,6,7	125	500	75	750
F	6	3, 4, 5, 6, 7, 8	150	500	75	750

Chapter 10

Microphone Virtualization by Neural Networks

The observation filter estimation by means of LMS/RLS or MMSE algorithm is a simple yet effective approach, that is however intrinsically constrained by limitations: these methods assume that the primary paths, $P(z)$ and $\Pi(z)$ are linear, stationary channels. Moreover, the virtualization task is also hindered by the distance between monitoring and the virtual microphones.

For this reason, the use of a Neural Network (NN) may be considered as an alternative to the traditional methods that may be worth investigating. In fact, unlike linear filters, the main advantages of a NN is the possibility to model complex non-linear and non-stationary channels. Especially in the automotive environment, where the acoustic channels are strongly time-varying, e.g., due to the car speed, this capability may have the potential to improve microphones virtualization performance.

Thanks to the recent popularity of deep learning techniques in visual contexts, significant interest in the sound processing field toward machine learning frameworks has emerged [58, 59]. The Deep Neural Network (DNN) can be trained to perform complex operations through supervised learning [60]. The network adjusts a set of weights using the mechanism of back-propagation, in order to find solutions to problems which would be otherwise hard to achieve, e.g., due to the difficulty of defining an accurate model. In the acoustic field, NNs have been successfully employed in several applications, such as acoustic modeling, speech and sound event recognition and music

classification. Motivated by the success of these works, a deep neural network based on the TCN architecture [61], trained for the purpose of multi-channel microphone virtualization is discussed. The work here presented can be regarded as a preliminary investigation on the subject can be found in [62].

This chapter is organized as follows. In Section 10.1, the main fundamentals on NNs are introduced. In Section 10.2, the considered TCN architecture for microphone virtualization is presented. Finally, Section 10.4 presents the main obtained numerical results on microphone virtualization through TCN.

10.1 Neural Network Background

In general, NN are compositions of units called neurons, denoted as ϱ, which get activated through connections weighted by parameters w, usually called weights. The activations are also conditioned on non-linear functions denoted as ς which act as gates, letting information through depending on the input value.

Non-linear functions are used to "activate" certain neurons which are associated with meaningful patterns based on which the network produces its results. Neurons can be disposed in layers l, allowing a network to process iteratively the information provided by previous neurons for increasing depth. The combination of an appropriate set of weights and activations in a sufficient amount of layers allows the NN to detect certain patterns in the input data, which are used to produce the target output. As example, in Fig. 10.1, a three-layer neural network in shown, in which the output of the j-th neuron of the ℓ-th layer, i.e., ϱ_j^ℓ can be written as

$$\varrho_j^\ell = \varsigma\left(\sum_{i=1}^{\iota^{(\ell-1)}} w_{ij}^\ell \varrho_i^{(\ell-1)} + b_j^\ell\right) \tag{10.1}$$

where $\iota^{(\ell-1)}$ denotes the total number of neurons in layer $(\ell-1)$, ς identifies the non-linear activation function, w_{ij}^ℓ represents the weight associated to the connection between the $\varrho_i^{(\ell-1)}$ to the neuron ϱ_j^ℓ and b_j^ℓ denotes the bias term. The network is trained to optimize the coefficients w_{ij}^ℓ and b_j^ℓ for all the connected neurons at every layer.

Several functions can be used for activation and the most relevant are:

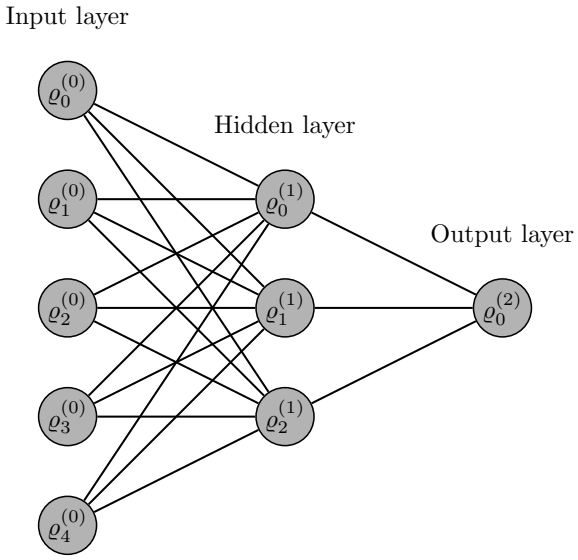


Figure 10.1: Example of neural network with fully connected layers. Each neuron ϱ_j^ℓ is the result of weighted connections from all the neurons of the previous layer.

- Sigmoid function: $\mathcal{S}(x) = \frac{1}{e^{-x}+1}$
- Hyperbolic tangent function: $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$
- Rectified Linear Unit (ReLU): $\text{ReLU}(x) = \max(0, x)$
- leaky ReLU: $\text{leakyReLU}(x, 0.01) = \max(0.01x, x)$

Each activation function presents different advantages, but generally the reasoning behind the choice of which one to use is based on the measured performance of the network. For instance, the ReLU activation function gained popularity as it showed great performance in the training of very deep networks due to its resistance to the problem of vanishing gradients [63].

In order to properly tune the parameters of a network, the optimal route in supervised learning (training a neural network employing references of the target output data) is to proceed through the principle of gradient descent. An error function is a measure of the distance between the true value and its predicted one, differentiable with respect to its inputs. The error function is applied at the output of the neural network and its gradient is computed with respect to all the parameters in the network. The aim of a training procedure is to adjust these parameters in order to minimize the

error function for all the possible inputs, which means searching its global minimum point in the space of all the parameters. This yields a problem of gradient descent in $\mathcal{P}$-dimension, where $\mathcal{P}$ is the total number of parameters. The learning rate of the algorithm has to be regulated in order to avoid that the algorithm gets stuck in local minimum points. Traditionally, the calculation of derivatives can be carried out in two was: by numerical differentiation or by symbolic differentiation. The alternative, deriving from a more general family of techniques called Automatic Differentiation (AD), is represented by the back-propagation algorithm, which is a more efficient method to calculate derivatives [64].

Mini-batch training is an optimization procedure for the training process of NNs and consists in using a group (batch) of input data samples for a single training iteration, instead of employing individual samples in separate iterations. The main motivation behind this procedure is that reducing the number of iterations may speed up the training procedure. The procedure works correctly if the input data has elements similar enough to each other to bring the same contributions to the gradients during training: under this condition, using mini-batches accelerates the convergence to the desired set of parameters without degrading the resulting network performance [64].

In 1980, in the image processing scenario, the concept of Neocognitron was introduced [65], yielding the so-called Convolutional Neural Network (CNN). Instead of using fully connected layers in which each neuron processes a single input image pixel, a 2-dimensional convolutional layer of weights was introduced. Convolutional layers apply the convolution operation between the input data and a filter of trainable weights. The discrete convolution operation for 2-dimensional arrays is defined as

$$y[a,b] = x[a,b] \otimes h[a,b] = \sum_{\ell=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x[\ell,n] h[a-\ell, b-n]$$

where $x[a,b]$ is the input, $h[a,b]$ denotes the trainable filter and $y[a,b]$ the output. The use of convolutional layers in deeper networks and the use of Graphical Processing Unit (GPU) to perform very fast matrix multiplications, ultimately allowed CNN to become the best performing and most popular solution for image classification in recent years [66].

Convolutional layers are generally defined by their filter size and their number of

feature maps [67]. A feature map represents a significant feature extracted through a dedicated filter, meaning that for a specified number of feature maps, the network will train an equal number of sets of weights. These feature maps are then combined through further convolutions, to learn more complex representations of the data. For a sufficiently deep network, after a sufficient number of training cycles, the neurons will learn to produce a significant representation of the input data through the feature maps.

10.2 TCN Architecture

TCNs were first introduced by Lea et al. in 2016 [68] and perform convolutions in the time domain. The TCN architecture is considered to be a Fully Convolutional Network (FCN) and typically are composed of several convolutional layers and a final dense layer. When FCNs were introduced in the image analysis and processing domain, the aim was to perform pixel-wise semantic segmentation. Contrarily to traditional CNNs, in which the size of the dense layers imposes limitations on the input picture, FCNs perform the classification task pixel by pixel without fixed constraints on the picture size.

Vanishing Gradients Vanishing gradients are a problem that affects the training of neural networks. Several effective solutions were proposed to attenuate their negative effects [69, 70]. During the training period, the gradients of the error function with respect to all the network weights are calculated through back-propagation, starting from the output and proceeding backwards. The problem of vanishing gradients arises when the calculated derivatives are very small, since the resulting gradient may become very close to zero as it back-propagates. Since the amount by which a weight is adjusted during back-propagation is proportional to the value of the derivative, vanishing gradients may slow down or even block the process of weight optimization of layers located at superficial levels, saturating thus the learning process earlier than convergence to an acceptable set of weights is achieved. Non-linear activations are the main cause of vanishing gradients since they present intervals where their derivatives are very close to 0.

As an example, let us consider the sigmoid activation function $\varsigma(x) = 1/e^{-x} + 1$

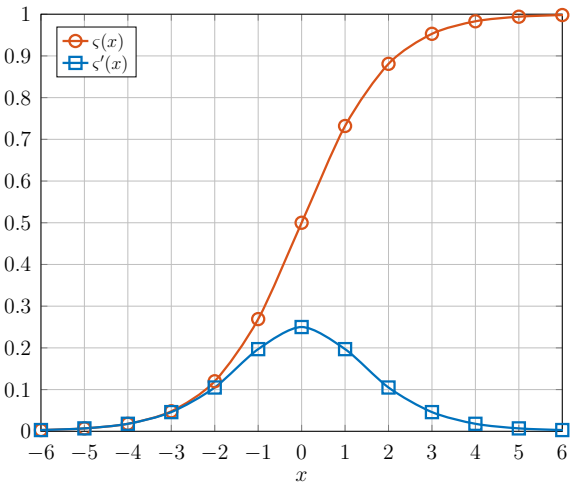


Figure 10.2: Example of vanishing gradients effect in case of sigmoid activation function.

whose derivative is

$$\varsigma'(x) = \frac{e^{-x}}{(e^{-x}+1)^2} \ .$$

In Fig. 10.2, the sigmoid activation function $\varsigma(x)$ and its gradient $\varsigma'(x)$ are shown as a function of the variable x. It is possible to observe that when the sigmoid reaches the saturation level, its gradient tends to vanish.

Residual Blocks The concept of residual blocks was introduced in CNN image classifiers to counteract the phenomenon of vanishing gradients; in fact, the deeper the neural network, the larger the number of vanishing gradients events [71]. To cope with this issue, the so-called *residual blocks* can be introduced. A residual block can be described as a stuck of layers that are in built in such a way the output of one layer is added to another layer that is in the block of deeper layer.

Considering a dense layer as expressed in (10.1), the output of the layer with the residual blocks can be written as

$$\varrho_j^{\ell} = \varsigma\left(\sum_{i=1}^{\iota^{(\ell-1)}} w_{ij}^{\ell}\varrho_i^{(\ell-1)} + b_j^{\ell}\right) + \varrho_i^{(\ell-1)} \ .$$

An intuitive block diagram of a skip connection for a single neuron is shown in Fig. 10.3.

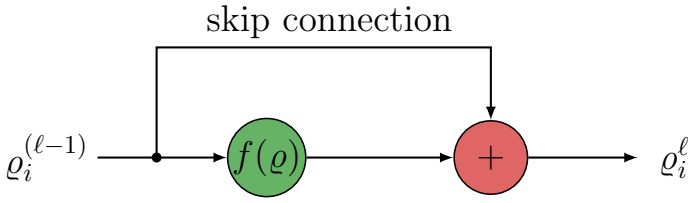


Figure 10.3: Skip connection for a single neuron.

Using a standard module of CNN, the so-called "skip connection" allows the gradient to back-propagate to lower layers through the addition of the input, that, since it is not conditioned on ϱ, it is not subject to the risk of vanishing.

Dilated Convolutions In convolutional networks, the receptive field r_f of a neuron can be defined as the region of the input space that affects the value of that neuron. In the scope of TCNs, the receptive field represents the memory of the input samples based on which the network produces its outcome. Hence, the larger the receptive fields the larger the memory of the network. Increasing the size of the convolutions allows to increase the size of the receptive fields, but it may incur to exhaustively long training time. TCNs, therefore, are based on small, increasingly dilated convolutions, operated at different layers in a pyramidal structure. The dilated convolution operation, in the 1D domain, can be defined as follows

$$y[n] = x[n] \circledast h[n] = \sum_{m=-\infty}^{+\infty} x[m]h[n - \vartheta m]$$

where $\circledast$ denotes the dilated convolution operator and ϑ is the dilation factor which depends on the network depth. Exploiting this mechanism, lower layers can capture the relationship between closer samples, while deeper layers can capture the relationship between farther samples.

Skip Connection Since the data extracted by the network is propagated from the bottom blocks to the top blocks of the pyramidal structure representing the TCN, there is the possibility that at higher layers the network drops some of the features learned at lower layers, which may contain relations between closer samples.

To ensure that all layers are considered equally, skip connections can be used instead

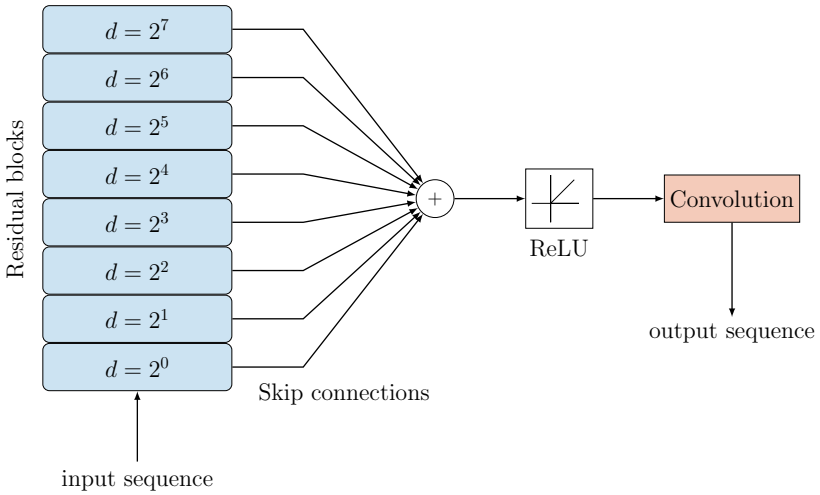


Figure 10.4: Schematic example representing a TCN using skip connections.

of extracting the output from the top layer. These connections are a secondary output of each residual block, and they are used to allow each layer to provide the extracted features without losing significance and combine them in a unified feature map, based on which the output sequence is produced. Notice that these are different from the residual connections, which are instead used for the purpose of bypassing the non-linearities in back-propagation to avoid the problem of vanishing gradients. In Fig. 10.4, skip connections are used to combine the output of all the 8 residual blocks in a single feature map, which is then passed through a ReLU activation and a convolution block to produce the output sequence.

10.3 Considered TCN Model

In this section, a description of the model considered in this work is provided.

The residual block model describes the core operations performed by the considered TCN at a layer level. Figure 10.5 represents a residual block at layer ℓ. The input of the residual block is the output of the previous block (at layer $\ell - 1$). The input of the first residual block (at layer $\ell = 1$) is directly the input data. The block is composed of two dilated convolutions, namely conv_A^ℓ and conv_B^ℓ, with dilation factor $\vartheta = 2^{\ell-1}$, with different parameters.

The weights associated with conv_A at layer ℓ are denoted as $w_{Akj}[n]$, where n is

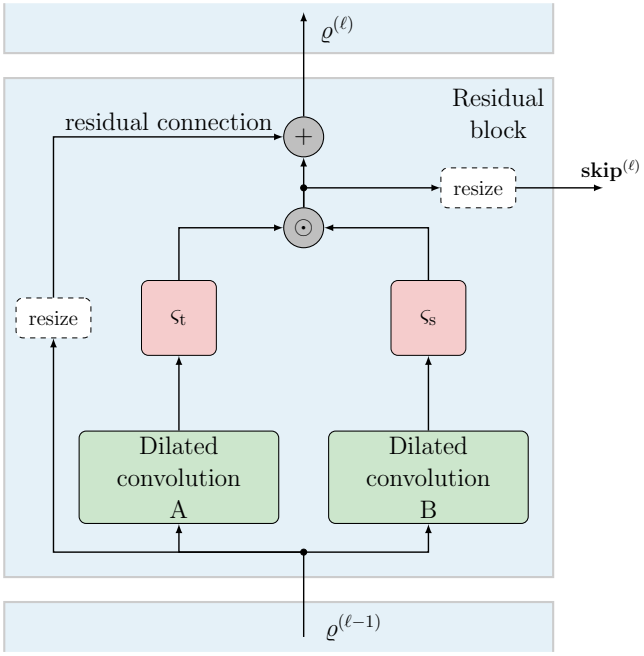


Figure 10.5: Representative scheme of the residual block employed in the considered model.

the temporal index (along the filter size dimension), k represents the associated input channel and j identifies the associated feature map. A similar procedure is used for conv_B by substituting the index B in place of A. At the output one obtains

$$\mathsf{conv}_A\left(\varrho^{(\ell-1)}[n]\right) = \sum_{k=1}^{K} w^{\ell}_{Akj}[n] \circledast \varrho^{(\ell-1)}_k[n] + b_{Aj}[n]$$
$$\mathsf{conv}_B\left(\varrho^{(\ell-1)}[n]\right) = \sum_{k=1}^{K} w^{\ell}_{Bkj}[n] \circledast \varrho^{(\ell-1)}_k[n] + b_{Bj}[n] \ .$$

These convolutions are followed by a gated activation function, which is obtained as the term-by-term product of the two non-linearities $\varsigma_t(x) = \tanh(x)$ and $\varsigma_s(x) = \mathcal{S}(x)$ applied to the results of the conv_A and conv_B operations, respectively. The combination of the convolutions and gated activation is therefore denoted by the residual signal r^{ℓ},

$$r^{\ell} = \varsigma_t\left(\mathsf{conv}_A\left(\varrho^{(\ell-1)}[n]\right)\right) \odot \varsigma_s\left(\mathsf{conv}_B\left(\varrho^{(\ell-1)}[n]\right)\right)$$

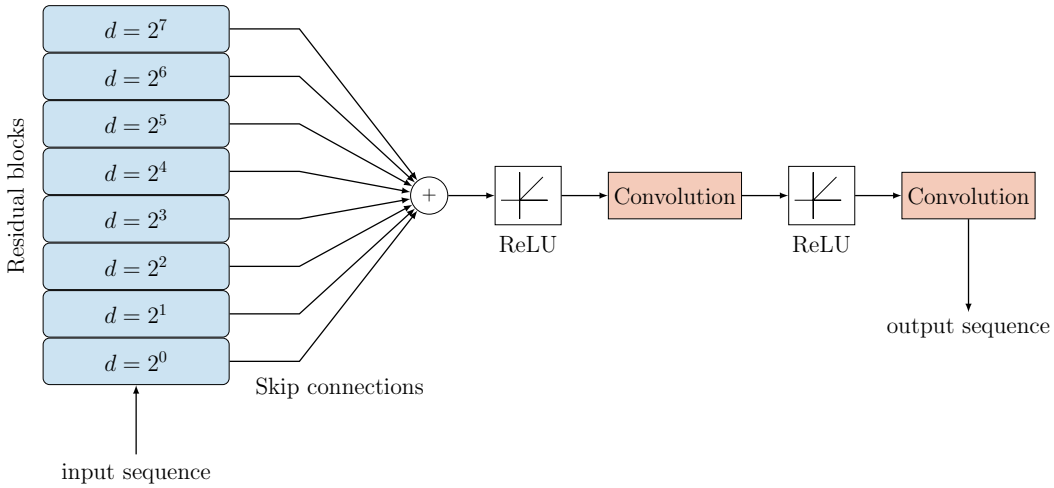


Figure 10.6: Schematic example representing the considered TCN model using skip connections.

where $\odot$ denotes the element-wise multiplication operator.

At this point, the input $\varrho^{(\ell-1)}$ is added to r^ℓ. Since dimensional mismatch issues may be occur, an optional resizing convolution operation (conv_r) is performed, yielding

$$\varrho_j^\ell = \mathsf{conv}_r \left(\varrho_j^{(\ell-1)} \right) + r^\ell \ .$$

The same kind of resizing convolution can be used at the **skip connection** which propagates the residual information at the output of the block for the same reason: residuals extracted at different layers may have a different dimensionality if the feature size is set to be layer-dependent. In this work the feature size at each layer is kept constant.

As previously mentioned in Sec. 10.2, skip connections carry out the features extracted from the residual blocks and average them together into a single feature map.

In Fig. 10.6, the considered system is shows. The sum is followed by two units composed of a ReLU activation function and a convolution. The first unit is used to expand the feature map into a feature map of higher dimension in order to train the system to recognize a larger number of features, whereas the second unit is used to compress the feature map to fit the dimensions of the desired output, that is the number of target virtual microphones V. These convolutions are not dilated.

The TCN model just described was implemented through the MATLAB® pro-

gramming language version 2020b, with the following additional Toolboxes:

- Audio Toolbox Version 2.3
- Communications Toolbox Version 7.4
- DSP System Toolbox Version 9.11
- Deep Learning Toolbox Version 14.1
- Parallel Computing Toolbox Version 7.3
- Signal Processing Toolbox Version 8.5
- Statistics and Machine Learning Toolbox Version 12.0.

The model hyperparameters of the network can be defined as the variables that control the learning process. In the considered model, the most relevant hyperparameters are:

- Number of TCN layers
- Feature size
- Total number of training epochs
- Initial learning rate
- Learning rate drop factor
- Filter sizes in the convolution operations.

In Tab. 10.1, the adopted hyperparameters of the considered TCN are summarized.

Table 10.1: Summary of the adopted hyperparameters of the TCN.

Parameter	Valued
Number of TCN layer	10
Feature size at each layer	100
Number of maximum training epoch	5
Initial learning rate	$5 \cdot 10^{-3}$
Learning drop factor	0.25
Filter size for dilated convolutions	2
Filter size for resizing convolutions	1
Expanding convolution feature size	256
Compressing convolution feature size	1

10.4 Obtained Results

Numerical results on microphone virtualization performed by using the TCN described in Section 10.3, are shown in this section. The input microphone signals are those of the sedan described in Section 9.3.1, yielding four different driving scenarios, namely, `Rough 40`,`Smooth 90`, `Rough Variable` and `Smooth Variable`.

A comparison is drawn between the roof monitoring microphone near the driver seat position and the corresponding driver headrest microphone. The result, in terms of MSE, is shown in Fig. 10.7. Intuitively, since the headrest monitoring microphone is located closer to the ear, it exhibits slightly better performance when employed for the virtualization task. For this reason, in the following tests, the headrest microphone is considered.

10.4.1 Single Training

The MSE of the simulation carried on in the four driving scenarios, training the models separately for each scenario, is shown in Fig. 10.8(a). The corresponding SPL analysis, for the `Smooth 90` driving scenario is presented in Fig. 10.8(b). In particular, for the sake of clarity, only the frequency range of the first 1000 Hz is shown. It is possible to observe that, around 150–300 Hz the TCN achieves good performance.

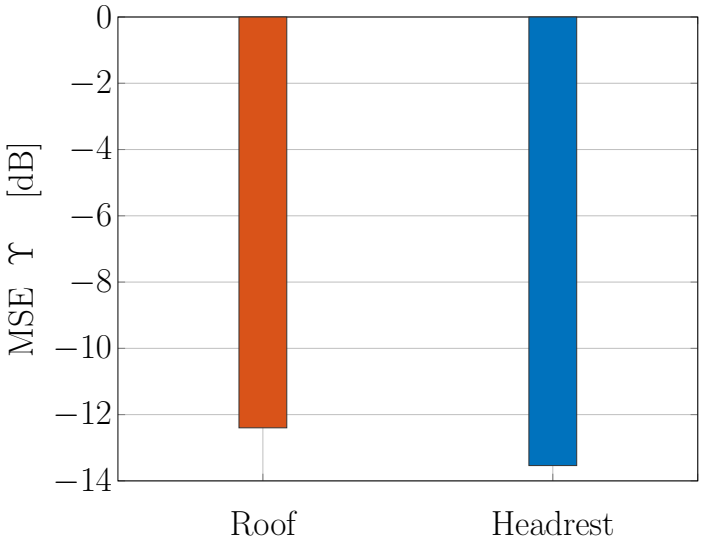


Figure 10.7: MSE performance for the TCN model in `Smooth 90` scenario, using the roof (orange) or the headrest (blue) monitoring microphone, with the models trained separately for each driving scenario.

10.4.2 Mixed Training

The results of the simulation carried on training the models with all the considered driving scenarios at once are shown in Fig. 10.9. In particular, the MSE analysis for all the driving scenarios and the SPL for the `Smooth 90` one only, are shown in Figs. 10.9(a) and (b), respectively. The results are approximately similar to the previous case, with the TCN introducing no improvements with respect to the linear model. Apparently, there were no significant improvements with the use of a TCN with respect to the case of individual trainings. Since the ASK data-set is very limited in size (the total duration is around 6 minutes and 40 seconds), it is possible that the network is not converging to a different result with respect to the linear filter because there is too little data for a sufficient training.

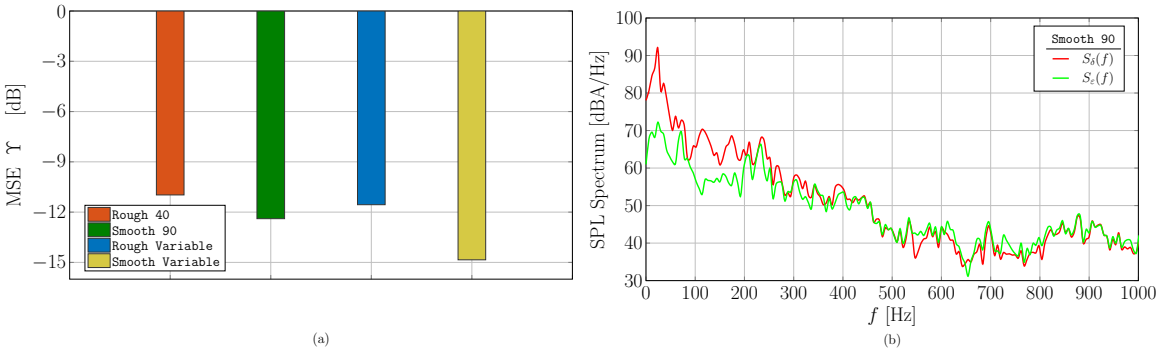


Figure 10.8: MSE analysis of TCN models with the sedan experimental measurements, with the models trained separately for each driving scenario.

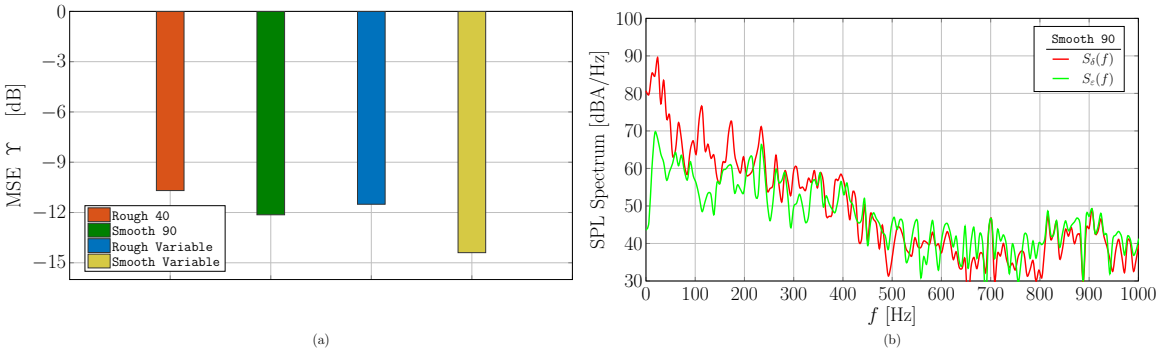


Figure 10.9: MSE analysis of TCN models with the sedan experimental measurements, with the models trained on all driving scenario.

Appendix on VMT

F Decomposition of Microphone Signals

Decomposition of microphone signals has been performed by using the MATLAB object octaveFilterBank. This function decomposes a signal into octave or fractional-octave sub-bands. In particular, a band is defined as an octave (1-octave sub-bands) when the highest frequency is two times the lowest one. Hence, it is possible to define a one-third ($^1/_3$-octave sub-bands) band when the upper frequency of the band equals the lower one multiplied by the cube root of two. Octave-band and fractional octave-band filters are commonly used to mimic how humans perceive sound. By considering the central frequency of an octave band f_c, the octave band edges $f_{\min}$ and $f_{\max}$ can be calculated by

$$f_c = \sqrt{2} f_{\min} = \frac{f_{\max}}{\sqrt{2}} \ .$$

In our MSE sub-band analysis, signals in (9.1), have been decomposed into 1- and $^1/_3$-octave bands by using filter banks as depicted in Fig. F.1. The sampling frequency is set to 6 kHz and the considered frequency range goes from 22 to 3000 Hz. The central frequencies f_c of these filters are shown in Tab. F.2

Table F.2: Central frequency f_c of filter bank for signal decomposition into 1- and $^1/_3$-octave bands.

Bandwidth	Central frequencies f_c [Hz]
1-octave band	31.63, 63.1, 125.9, 251.19, 501.19, 1000, 1995.27
$^1/_3$-octave band	25.12, 31.63, 39.82, 50.12, 63.1, 79.44, 100, 125.9, 158.49, 199.53, 251.19, 316.23, 398.11, 501.19, 630.96, 794.33, 1000, 1258.93, 1584.9, 1995.27, 2511.89

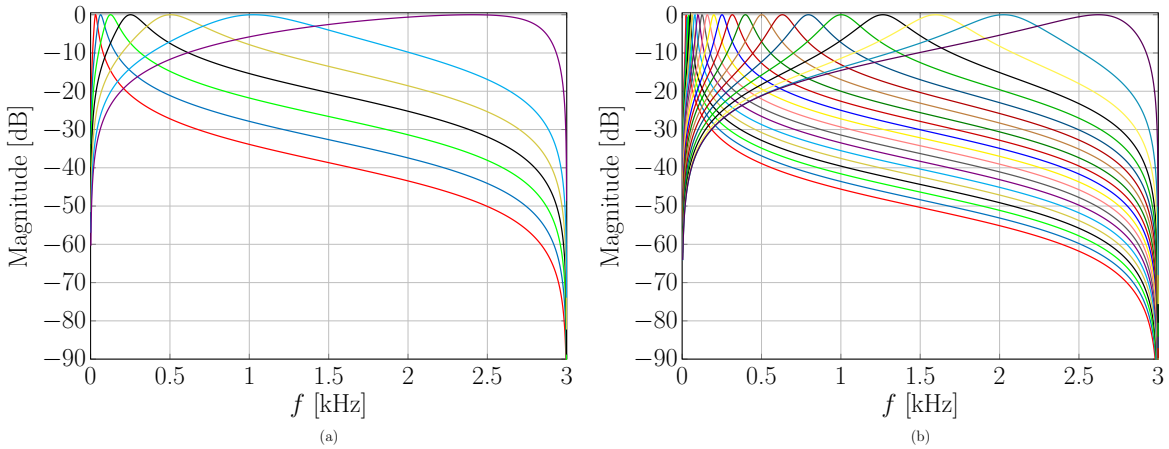


Figure F.1: Magnitude frequency response of filter bank for signal decomposition into 1-octave (a) and 1/3-octave bands (b).

G Further Results on VMT with Indirect Measurements

Further numerical results on the observation filter estimation performed by a single monitoring microphone to virtualize one quiet zone with indirect measurements (virtual microphone signal is synthesized starting from the monitoring one) are presented in this Appendix. In particular, results are organized based on the type of filtering that is applied to the monitoring microphone signal to obtain the virtual one, namely, synthetic and experimental responses, reported in Sections G.1 and G.2. More precisely, Section G.1 is divided into two subsections based on OF estimation approach, i.e., fixed and adaptive, in Sections G.1.1 and G.1.2, respectively. Further results on mid-size SUV and garage testing experimental responses are shown in Sections G.2.1 and G.2.2, respectively. Finally, in Section G.2.3, numerical results on the robustness under microphone mismatch for the case of garage testing experimental responses are shown.

G.1 Synthetic Responses

G.1.1 Fixed Approach: MMSE

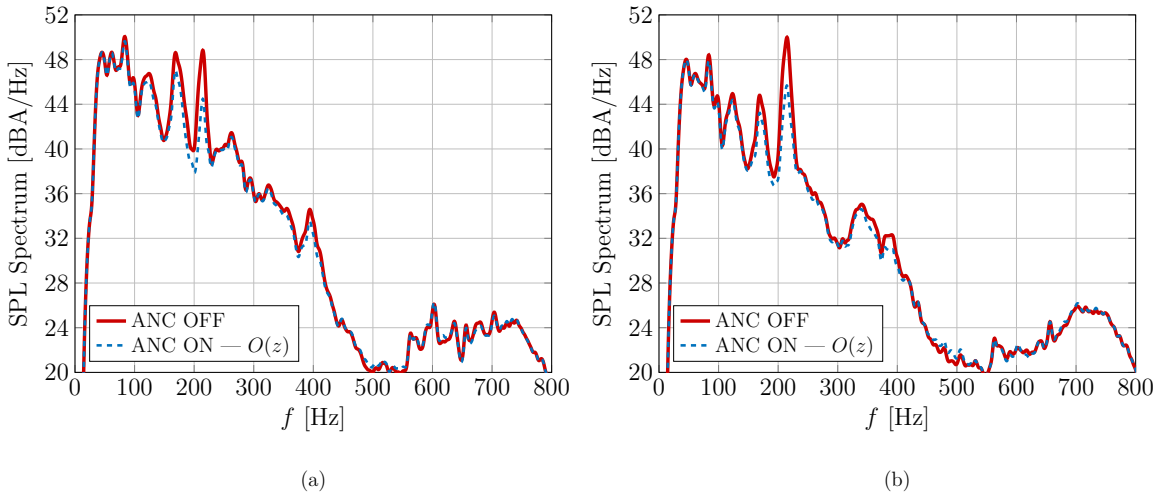


Figure G.1: SPL spectra for a multi-reference single-quiet-zone SISO system for ANC OFF and ANC ON with optimal and estimated observation filter for Experiment 4 and Scenario 2. Left (a) and right (b) virtual error microphones.

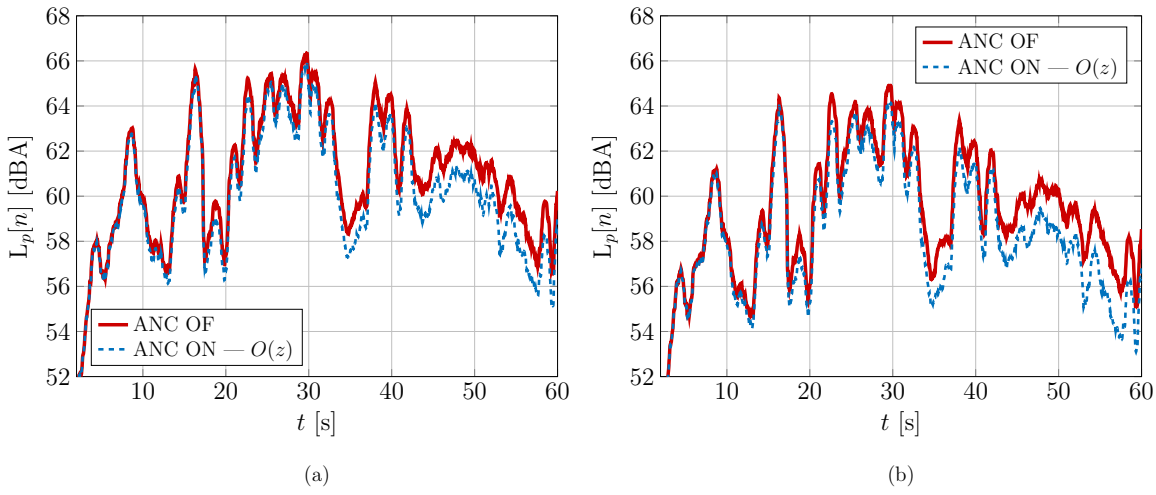


Figure G.2: Sliding window SPL against time for a multi-reference single-quiet-zone SISO system for ANC OFF and ANC ON with optimal and estimated observation filter for Experiment 4 and Scenario 2. Left (a) and right (b) virtual error microphones.

G.1.2 Adaptive Approaches: LMS & RLS

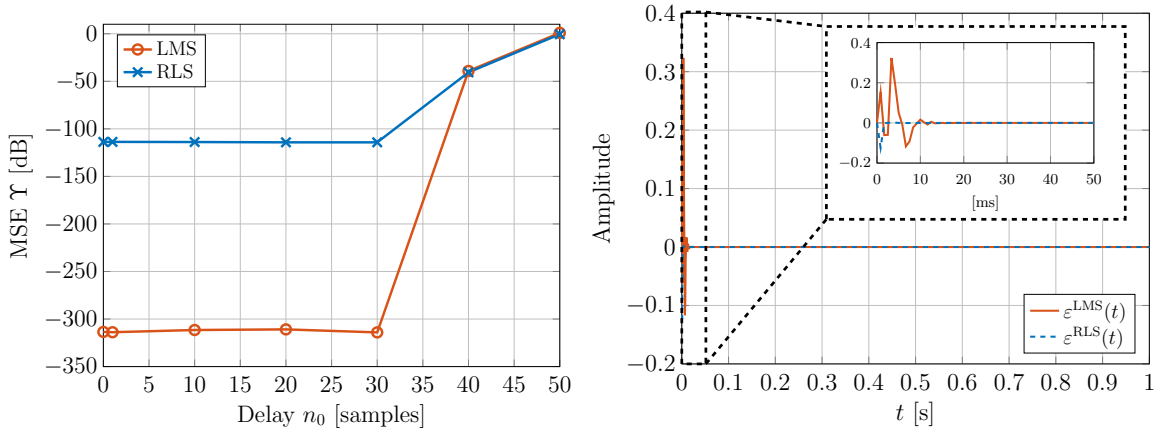


Figure G.3: MSE against delay n_0 in samples for LMS and RLS algorithms (a). Algorithm convergence speed against simulation time for the case of best delay $n_0^{\mathrm{opt}} = 0$ samples (b).

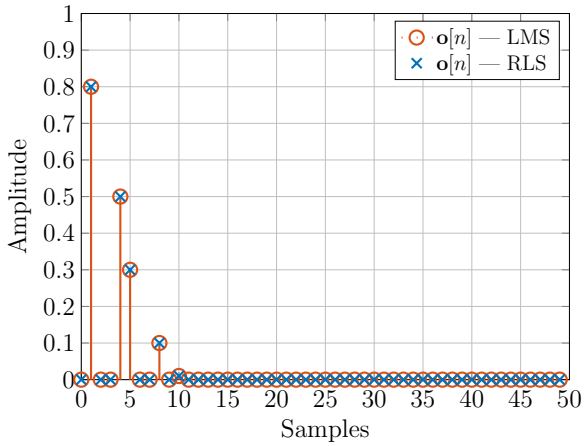


Figure G.4: Impulse responses of observation filter estimated by employing LMS and RLS algorithms for the case of best delay $n_0^{\mathrm{opt}} = 0$ samples.

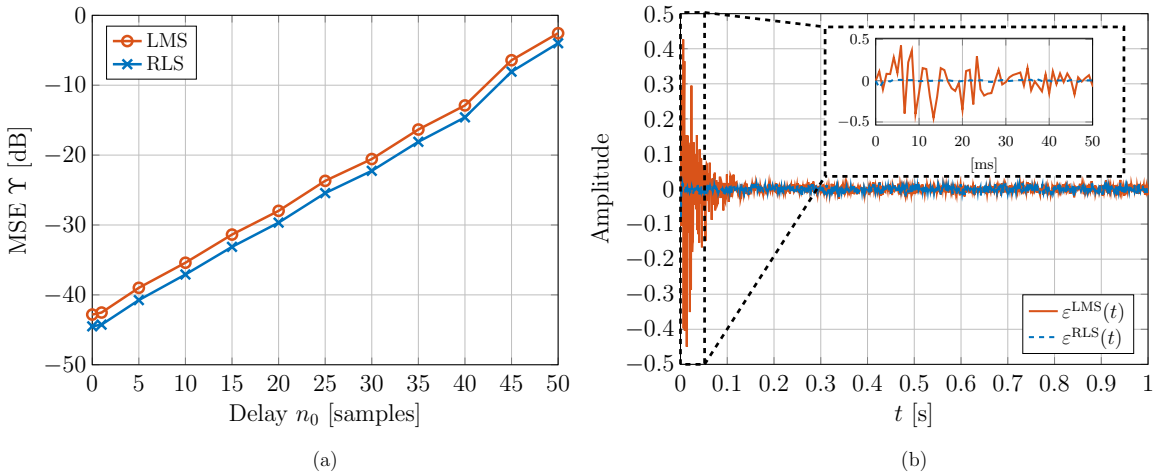


Figure G.5: MSE against delay n_0 in samples for LMS and RLS algorithms (a). Algorithm convergence speed against simulation time for the case of best delay $n_0^{\mathrm{opt}} = 0$ samples (b).

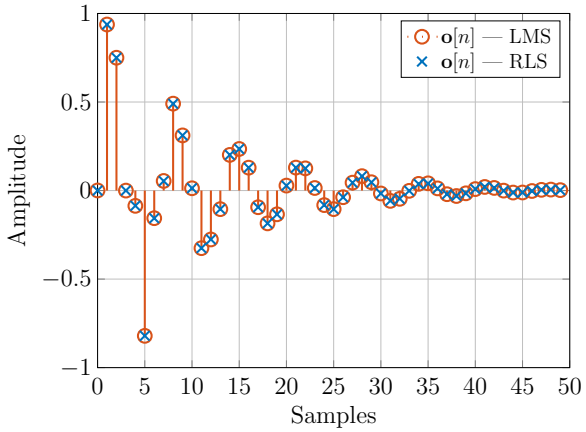


Figure G.6: Impulse responses of observation filter estimated by employing LMS and RLS algorithms for the case of best delay $n_0^{\mathrm{opt}} = 0$ samples.

G.2 Experimental Responses

G.2.1 Mid-size SUV

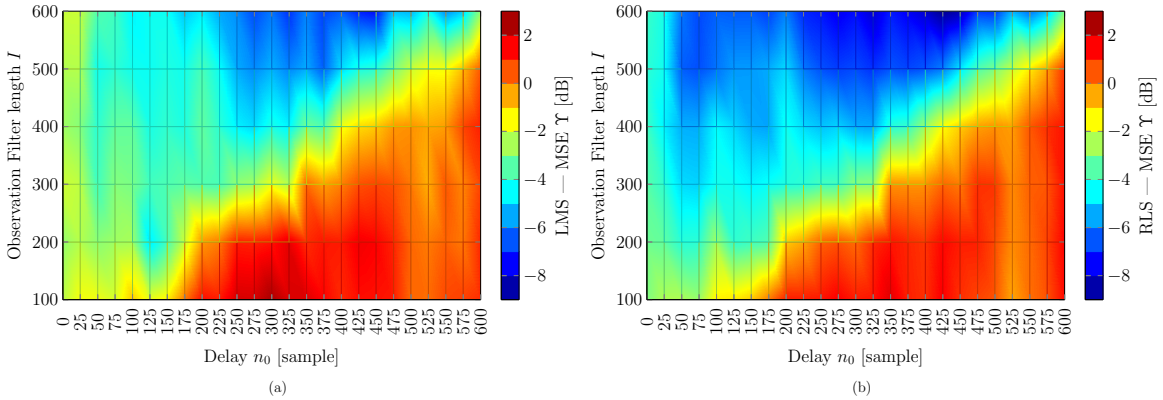


Figure G.7: MSE heat color map against observation filter length I and delay n_0 in samples for LMS (a) and RLS (b) algorithms for the Case 1 with mid-size SUV experimental responses.

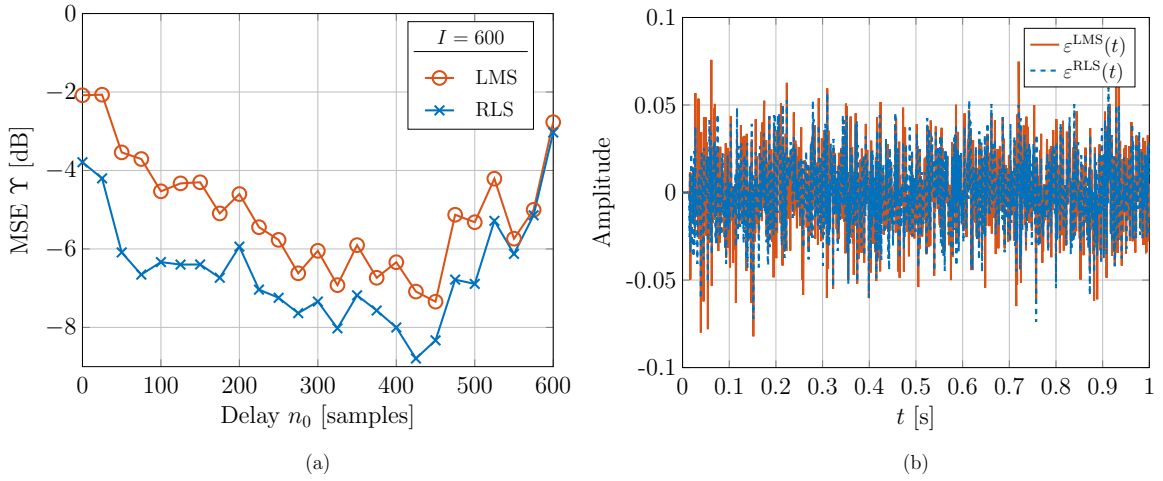


Figure G.8: MSE against delay n_0 in samples for LMS and RLS algorithms and observation filter length $I = 600$ (a). Algorithm convergence speed against simulation time for $I = 600$ and delay $n_0^{\mathrm{opt}} = 450$ samples (b). Case 1 with mid-size SUV experimental responses.

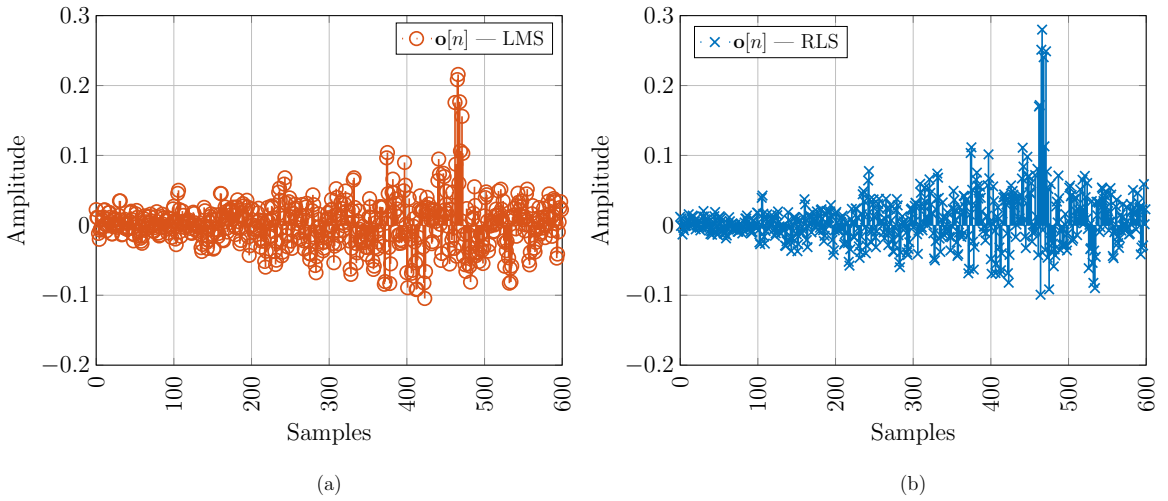


Figure G.9: Impulse responses of observation filter estimated by employing LMS (a) and RLS (b) algorithms for best delay $n_0^{\text{opt}} = 450$ samples. Case 1 with mid-size SUV experimental responses.

G.2.2 Garage Testing

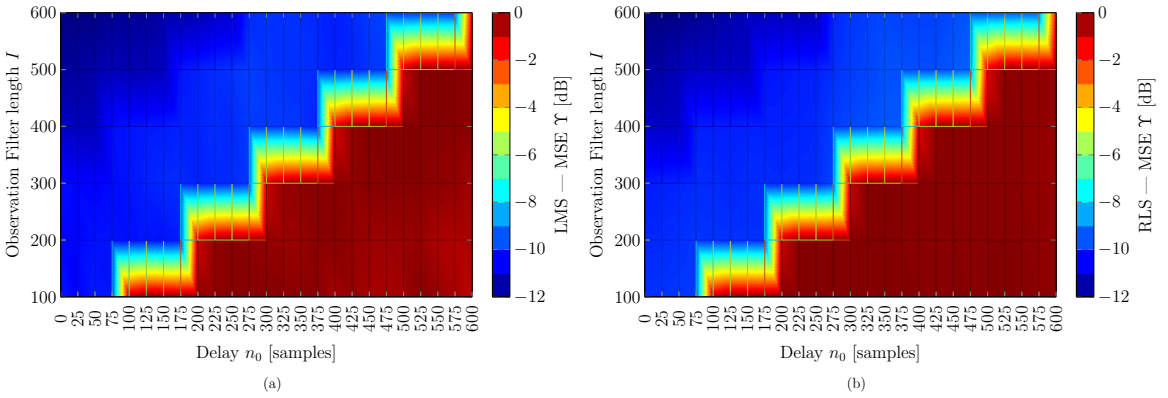
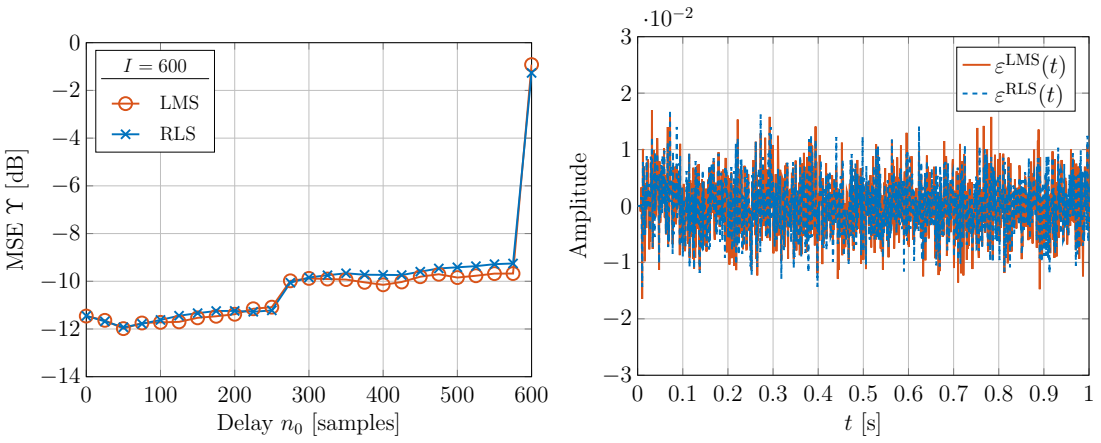
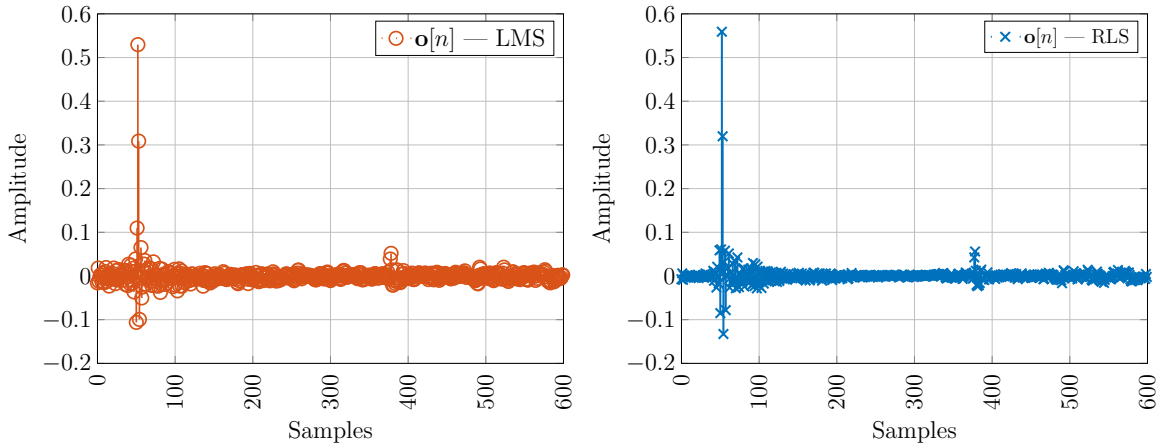


Figure G.10: MSE heat color map against observation filter length I and delay n_0 in samples for LMS (a) and RLS (b) algorithms for the Case 1 with garage testing experimental responses.



(a) (b)

Figure G.11: MSE against delay n_0 in samples for LMS and RLS algorithms and observation filter length $I = 600$ (a). Algorithm convergence speed against simulation time for $I = 600$ and delay $n_0^{\mathrm{opt}} = 25$ samples (b). Case 1 with garage testing experimental responses.



(a) (b)

Figure G.12: Impulse responses of observation filter estimated by employing LMS (a) and RLS (b) algorithms for the case of $I = 600$ samples and best delay $n_0^{\mathrm{opt}} = 25$ samples. Case 1 with garage testing experimental responses.

G.2.3 Microphone Mismatch (Garage Testing)

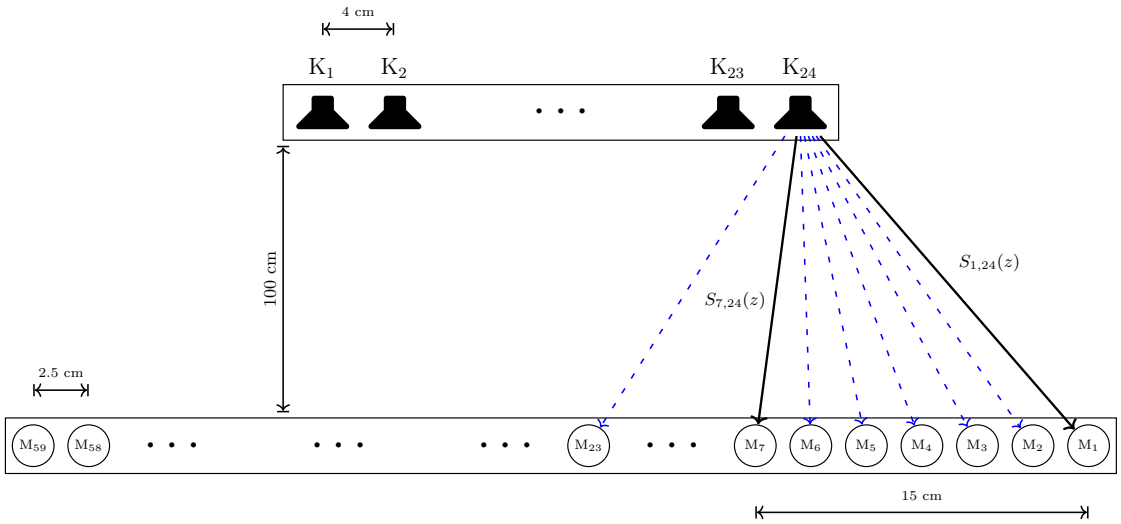
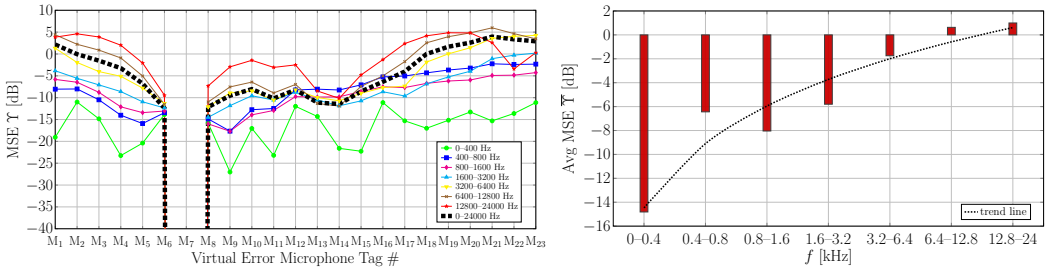


Figure G.13: Block diagram of garage testing experimental responses setup for OF robustness under virtual microphone mismatch.

Table G.3: Summary table of considered setup parameters for garage testing OF estimation robustness under virtual microphone mismatch.

Description	Value
Observation filter length	$I = 600$
Sampling frequency	$f_s = 48$ kHz
Speaker	K_{24}
Delay	$n_0 = 25$ samples
Primary path (for monit. mic.)	$P(z) = S_{1,24}(z)$
Monitoring microphone signal	$d[n] = u[n] \otimes p[n] = u[n] \otimes s_{1,24}[n]$
Primary path (for virt. mic.)	$\Pi(z) = S_{7,24}(z)$
Virtual microphone signal	$\delta[n] = u[n] \otimes \pi[n] = u[n] \otimes s_{7,24}[n]$
Primary path (for virt. mic. mismatch)	$\Pi'(z) = S_{i,24}(z)$, with $i = 1, 2, ..., 23$
Mismatched virtual microphone signal	$\delta'[n] = u[n] \otimes \pi'[n] = u[n] \otimes s_{i,1}[n]$



(a) (b)

Figure G.14: MSE Υ per sub-band against virtual error microphone tag number ($M_7 \to -\infty$ since $\delta'[n] = \delta[n]$) for the case of garage testing experimental responses (a). Average of Υ with respect to the virtual error microphones against sub-band (M_7 is not considered) for the case of garage testing experimental responses (b).

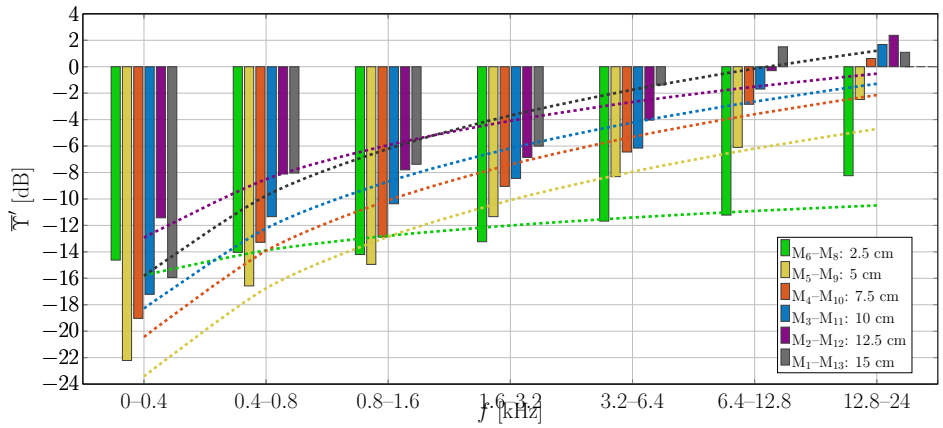


Figure G.15: Average of Υ with respect to equidistant virtual error microphones against sub-band for the case of garage testing experimental responses. Dashed curves denote corresponding trend lines.

H Further Results on VMT with Direct Measurements

In this appendix, further obtained results on the observation filters estimation performed by multiple monitoring microphones to virtualize multiple positions with direct measurements, are presented. This appendix is organized according to the considered experimental measurement, namely, sedan, luxury car and subcompact car in Sections H.1, H.2 and H.3, respectively.

H.1 Sedan

In this section, results on the observation filter estimation employing sedan measurements for the remaining driving scenarios (`Rough 40` and `Smooth Variable`) are considered. Observation filter estimation is performed by employing LMS and MMSE algorithm.

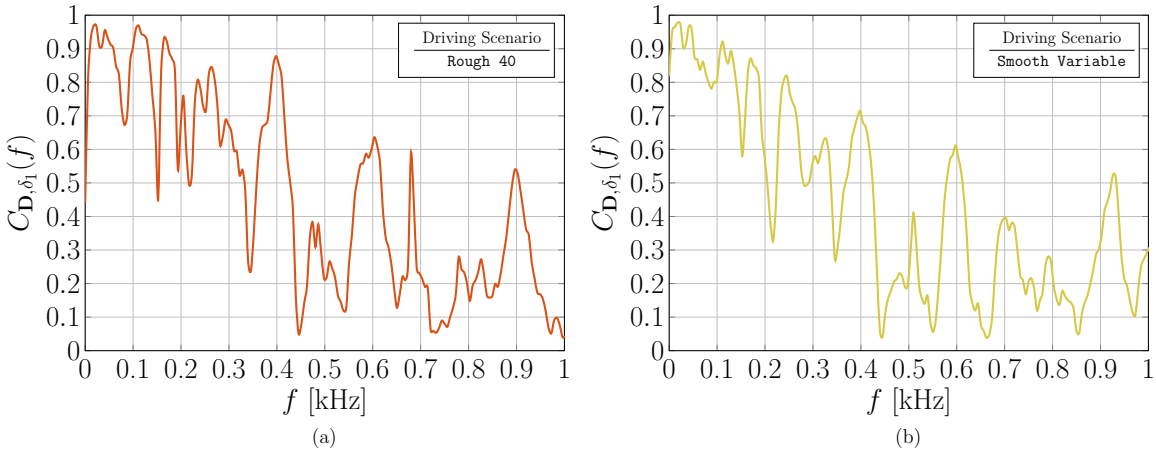


Figure H.1: Spectral coherence between head and roof monitoring microphone signals and the left virtual one for sedan measurements for `Rough 40` (a) and `Smooth Variable`(b) driving scenarios.

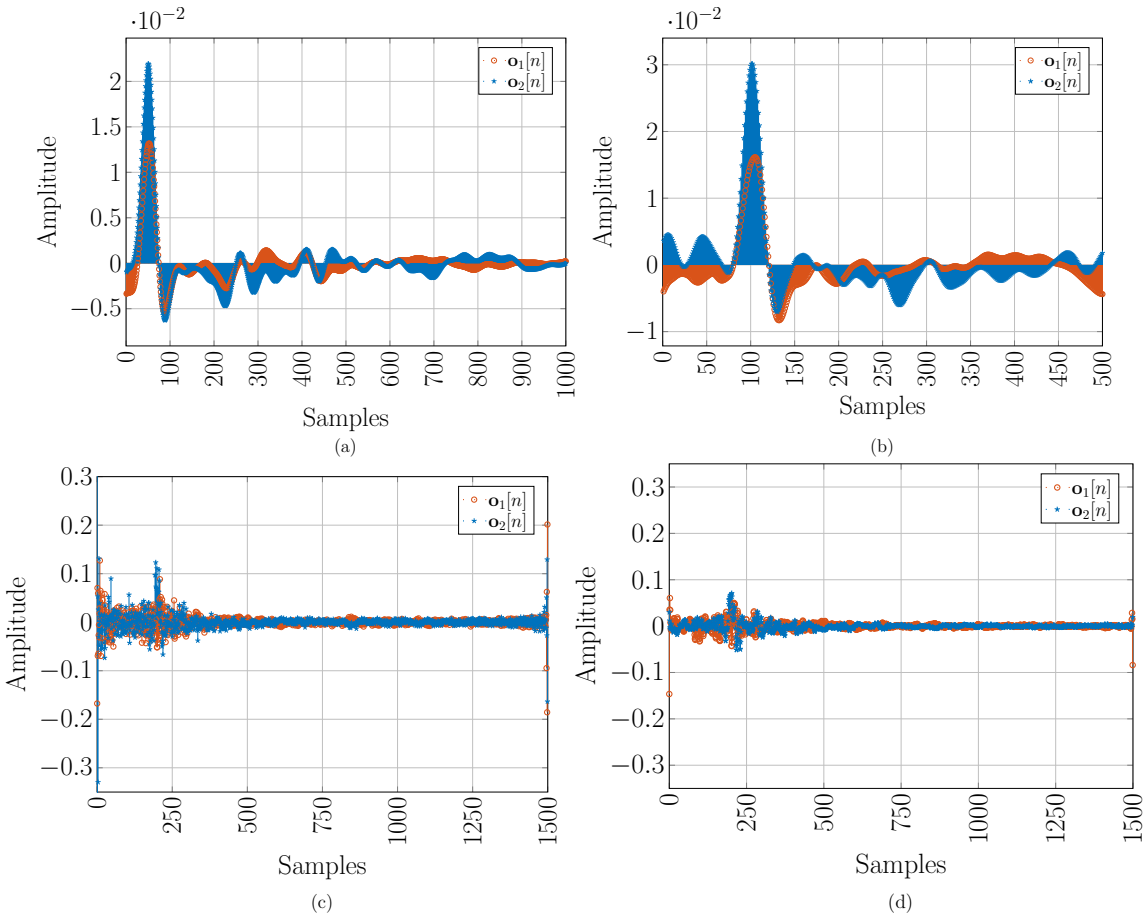


Figure H.2: Impulse response of estimated observation filter $O(z)$ for driver roof and driver head monitoring microphones and driver's left virtual microphone by means of the LMS algorithm for `Rough 40` (a) and `Smooth Variable` (b) and MMSE algorithm for `Rough 40` (c) and `Smooth Variable` (d) driving scenarios.

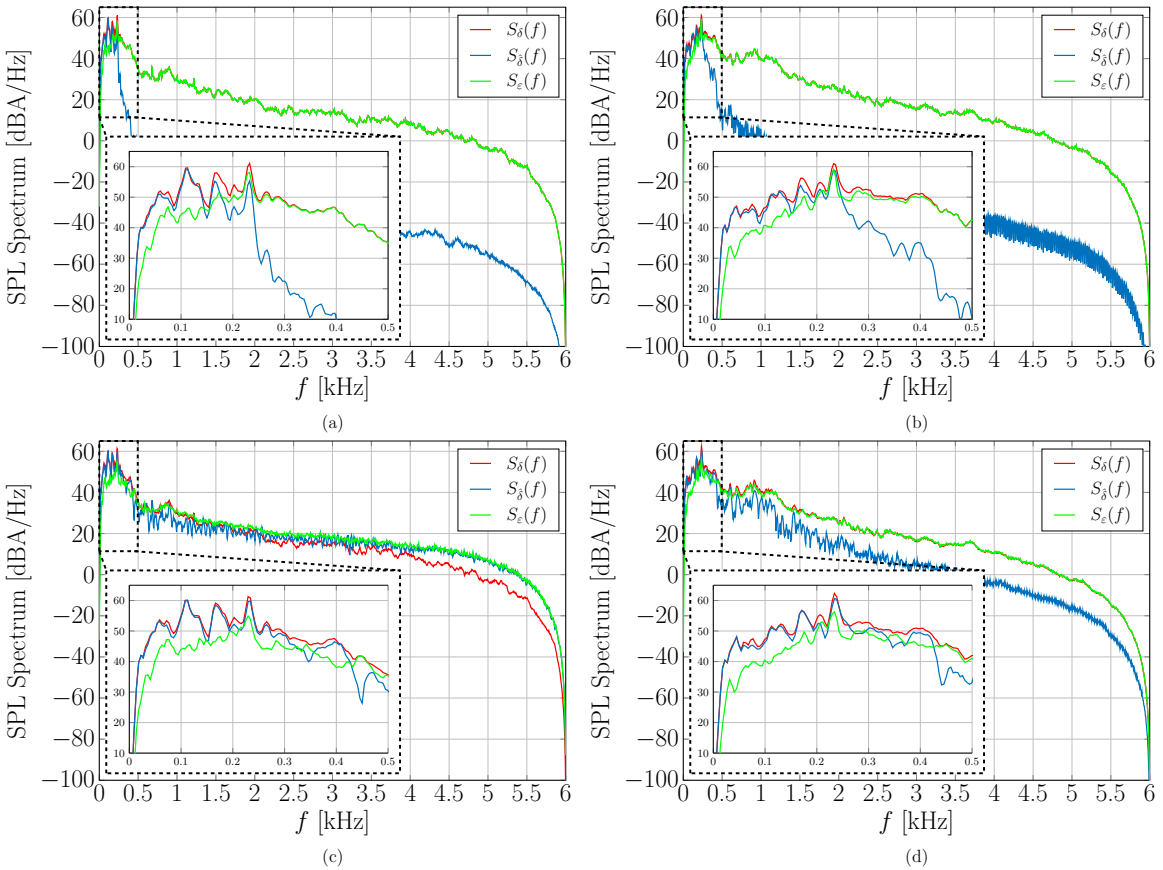


Figure H.3: SPL spectra of $\delta[n]$, $\hat{\delta}[n]$ and $\varepsilon[n]$ for driver roof and driver head monitoring microphones and driver's left microphone virtualization by means of the LMS algorithm for `Rough 40` (a) and `Smooth Variable` (b) and the MMSE algorithm for `Rough 40` (c) and `Smooth Variable` (d) driving scenarios.

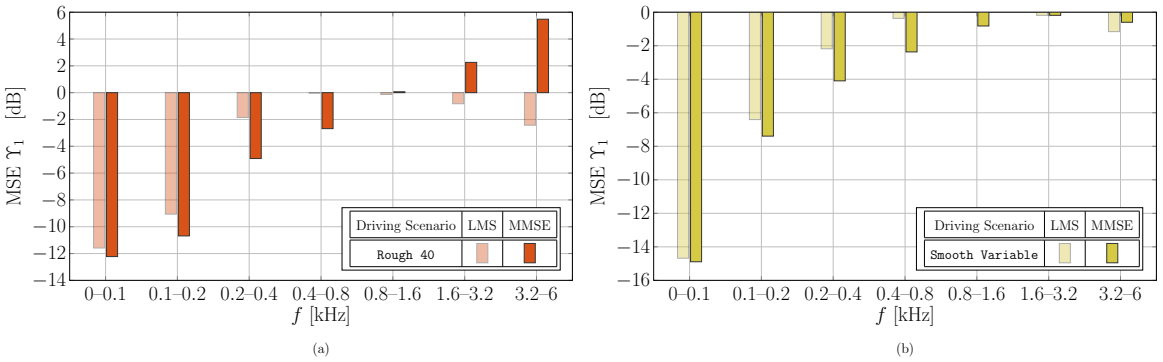


Figure H.4: MSE against frequency sub-band: comparison between observation filter estimation approaches for `Rough 40` (a) and `Smooth Variable` (b) driving scenarios.

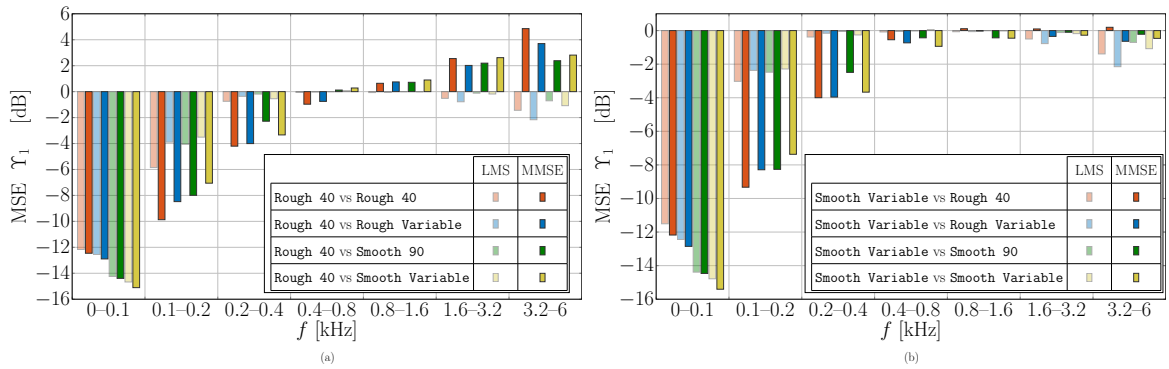


Figure H.5: MSE against frequency sub-band: comparison between observation filter estimation approaches calculated over the entire duration of `Rough 40` (a) and `Smooth Variable` (b) for driving scenario mismatch.

H.2 Luxury car

In this section, results on the observation filter estimation employing luxury car measurements for the remaining driving scenarios (`Rough` and `Traffic`) are considered in Sections H.2.1 and H.2.2, respectively. Observation filter estimation is performed by employing the LMS algorithm.

H.2.1 `Rough` Driving Scenario

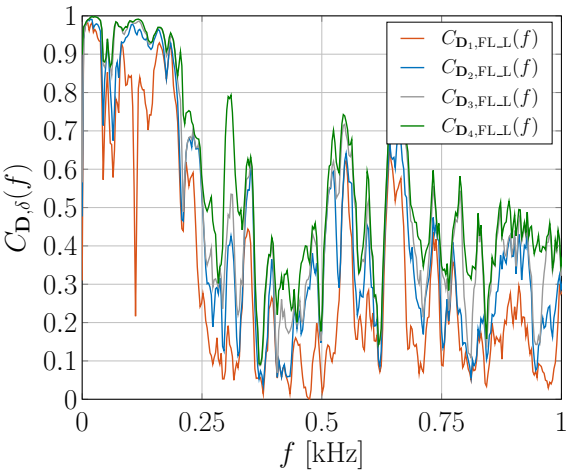


Figure H.6: Spectral coherence between monitoring microphone signals and the left virtual one for luxury car measurements for `Rough` driving scenario.

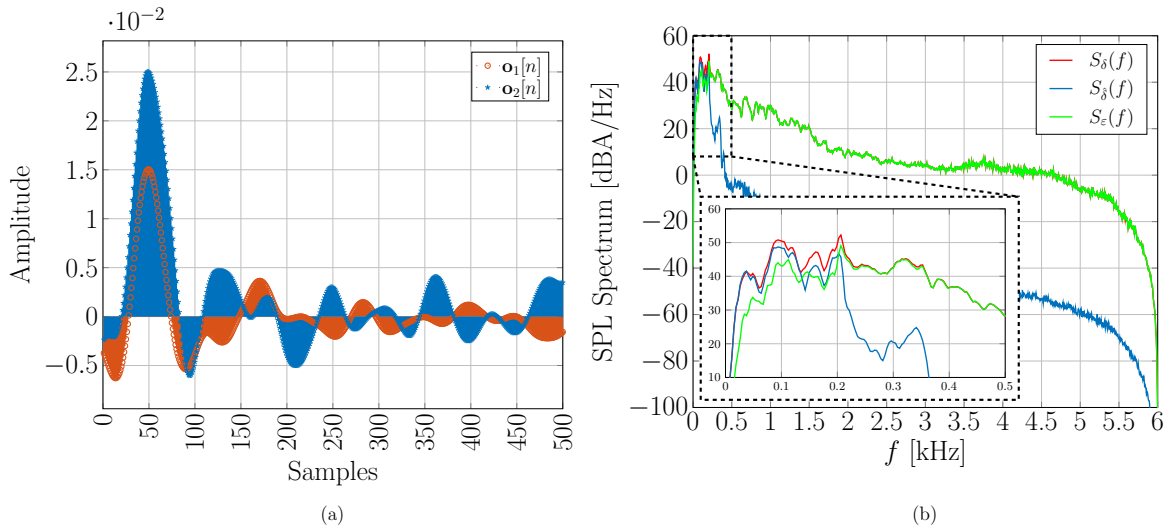


Figure H.7: Observation filter estimation results for luxury car experimental measurements for `Rough` driving scenario. Driver's left binaural microphone virtualization by means of monitoring microphone SQ3 and virtual microphone FL_R by employing LMS algorithm. Impulse responses of observation filters estimated for the case of best delay $n_0^{\text{opt.}} = 50$ (a). SPL spectra of target signal $\delta[n]$, its reconstructed version $\hat{\delta}[n]$ and their difference $\varepsilon[n]$ (b).

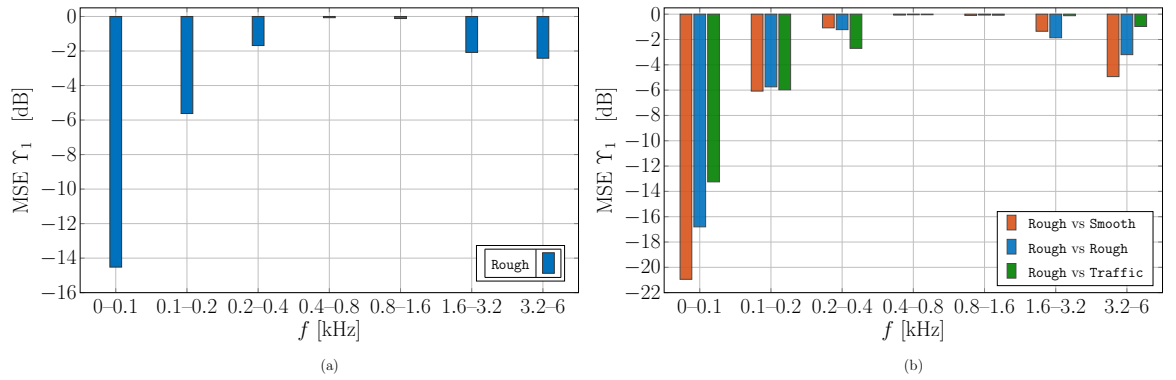


Figure H.8: MSE against frequency sub-band for luxury car experimental measurements for `Rough` driving scenario (a) and road mismatch (b).

H.2.2 `Traffic` Driving Scenario

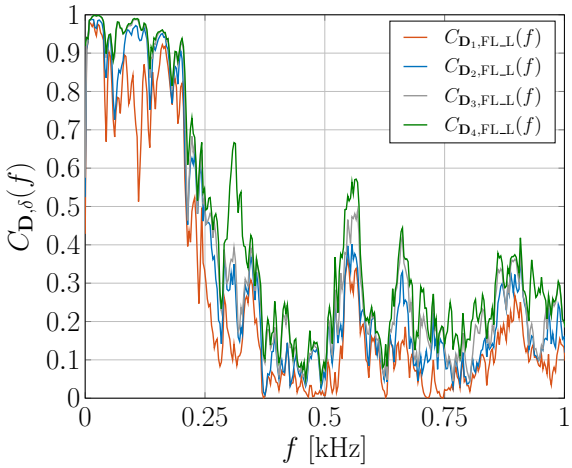


Figure H.9: Spectral coherence between monitoring microphone signals and the left virtual one for luxury car measurements for `Traffic` driving scenario.

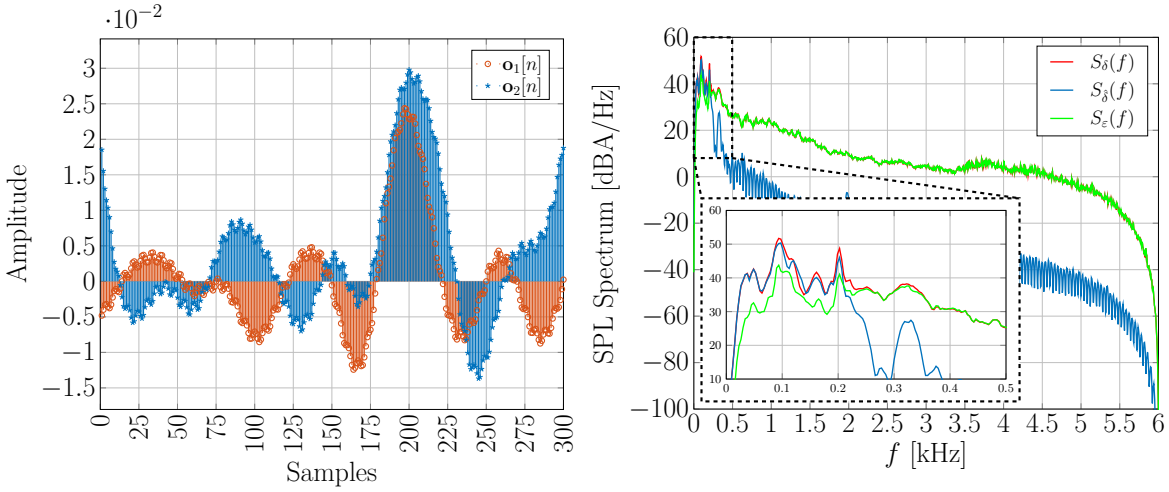


Figure H.10: Observation filter estimation results for luxury car experimental measurements for `Traffic` driving scenario. Driver's left binaural microphone virtualization by means of monitoring microphone SQ3 and virtual microphone FL_R by employing LMS algorithm. Impulse responses of observation filters estimated for the case of best delay $n_0^{\text{opt.}} = 200$ (a). SPL spectra of target signal $\delta[n]$, its reconstructed version $\hat{\delta}[n]$ and their difference $\varepsilon[n]$ (b).

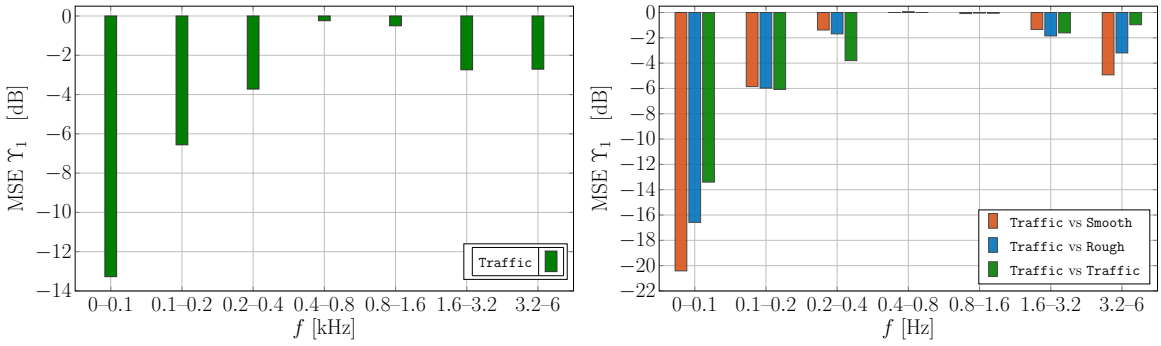


Figure H.11: MSE against frequency sub-band for luxury car experimental measurements for `Traffic` driving scenario (a) and for road mismatch (b).

H.3 subcompact car

In this section, results on the observation filter estimation employing subcompact car measurements for the remaining virtual microphone positions are presented. This section is organized according the considered setups, namely, `Roof` and `Headrest`, in Sections H.3.1 and H.3.2, respectively.

H.3.1 Roof Setup

The following results refer to the driver's right virtual microphone. Observation filter estimation is performed by employing LMS and MMSE algorithms.

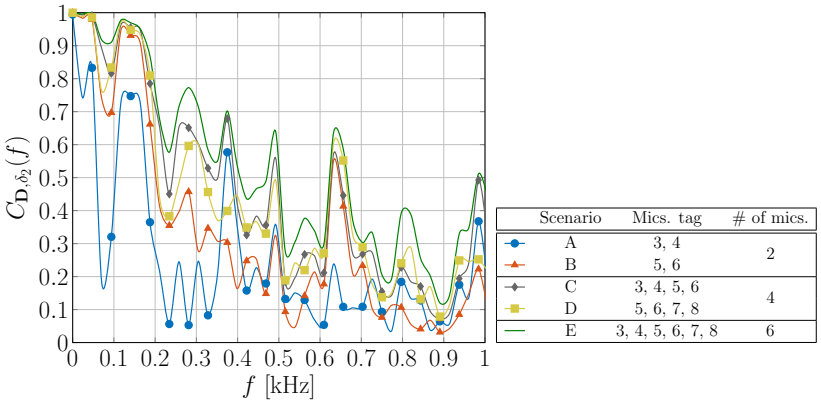


Figure H.12: Spectral coherence per scenario for the subcompact car experimental measurements and `Roof` setup between monitoring microphone signals and the driver's right virtual one.

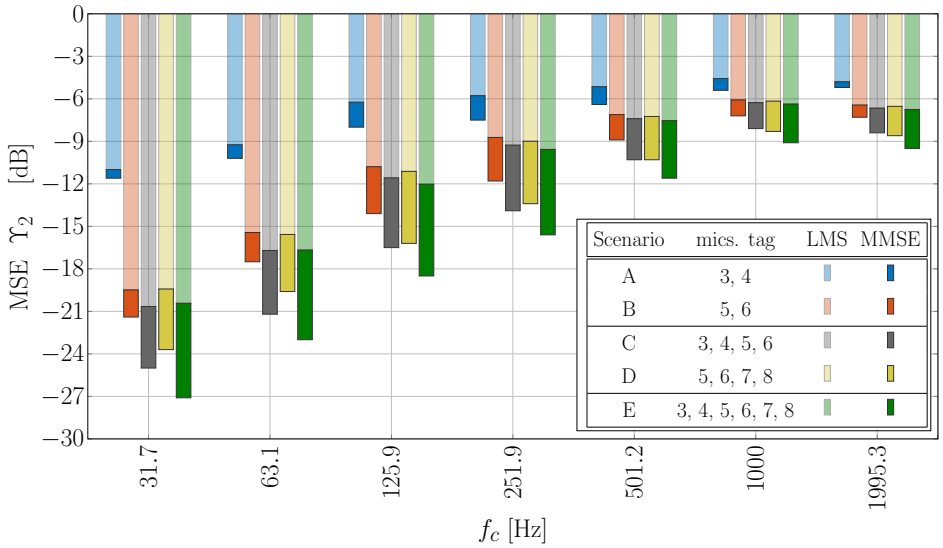


Figure H.13: Performance comparison in terms of normalized MSE as a function of octave band of adaptive (LMS) and fixed (MMSE) observation filter estimation approaches at the driver's right virtual microphone position for all the considered scenarios with subcompact car experimental measurements and `Roof` setup.

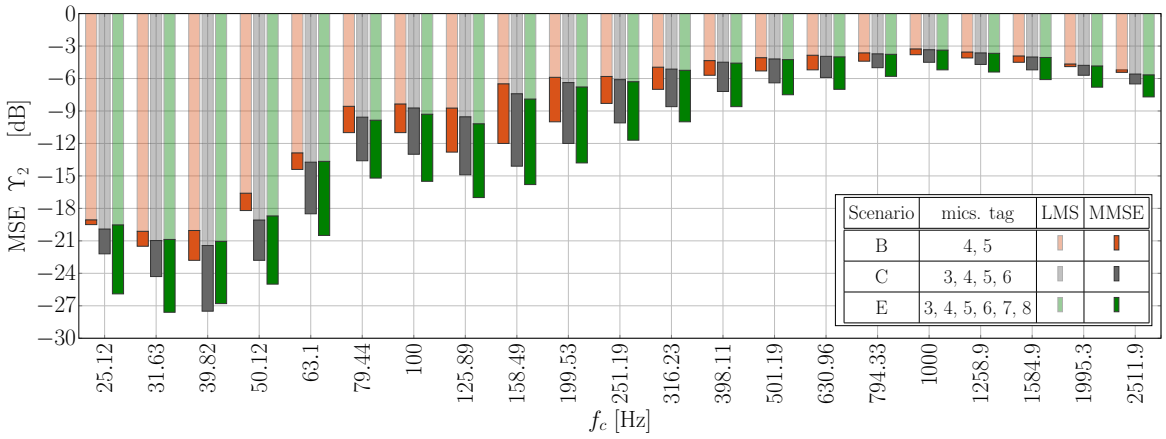


Figure H.14: Performance comparison in terms of normalized MSE as a function of 1/3-octave bands of adaptive (LMS) and fixed (MMSE) observation filter estimation approaches at the driver's right virtual position for the best scenarios, selected among those with the same number of monitoring microphones, with the subcompact car experimental measurements and `Roof` setup.

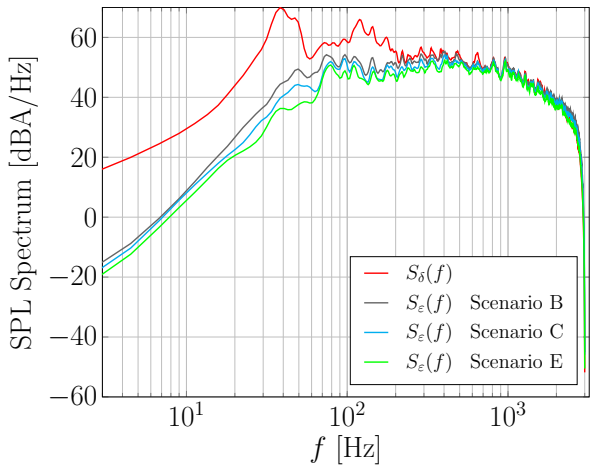


Figure H.15: SPL spectra of $\delta[n]$ and $\varepsilon[n]$ for the best scenarios, selected among those with the same number of monitoring microphones, for the driver's right microphone virtualization by means of the MMSE algorithm with subcompact car experimental measurements and `Roof` setup.

H.3.2 Headrest Setup

The following results refer to the passenger's left virtual microphone. Observation filter estimation is performed by employing LMS and MMSE algorithms.

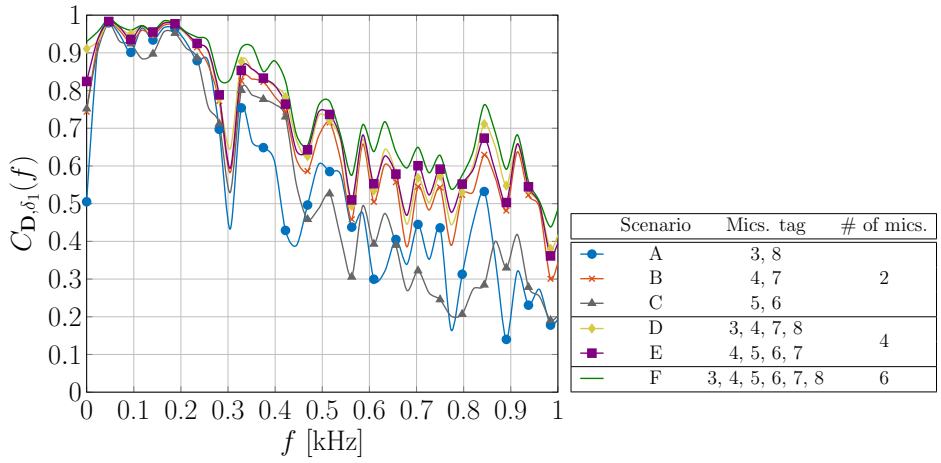


Figure H.16: Spectral coherence per scenario for the subcompact car experimental measurements and `Headrest` setup between monitoring microphone signals and the passenger's left virtual one.

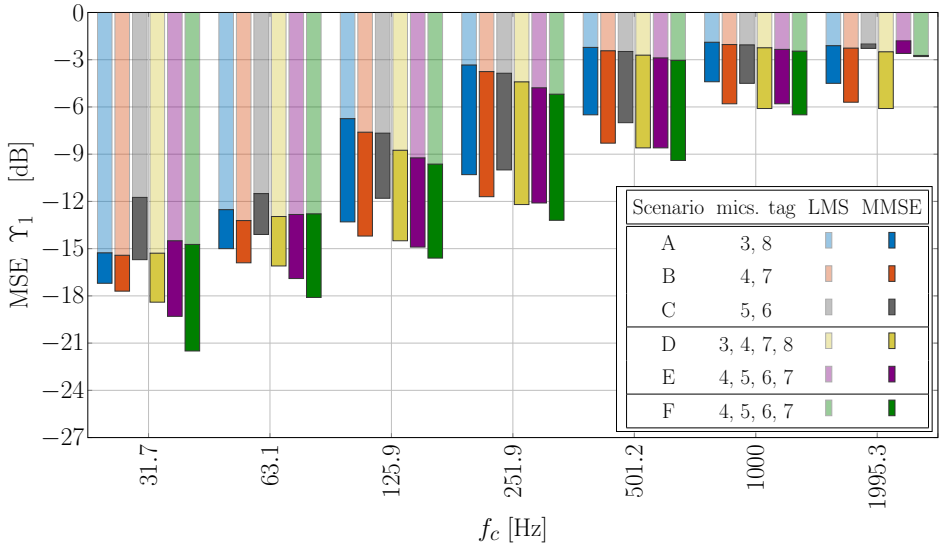


Figure H.17: Performance comparison in terms of normalized MSE as a function of octave band of adaptive (LMS) and fixed (MMSE) observation filter estimation approaches at the passenger's left virtual microphone position for all the considered scenarios with subcompact car experimental measurements and `Headrest` setup.

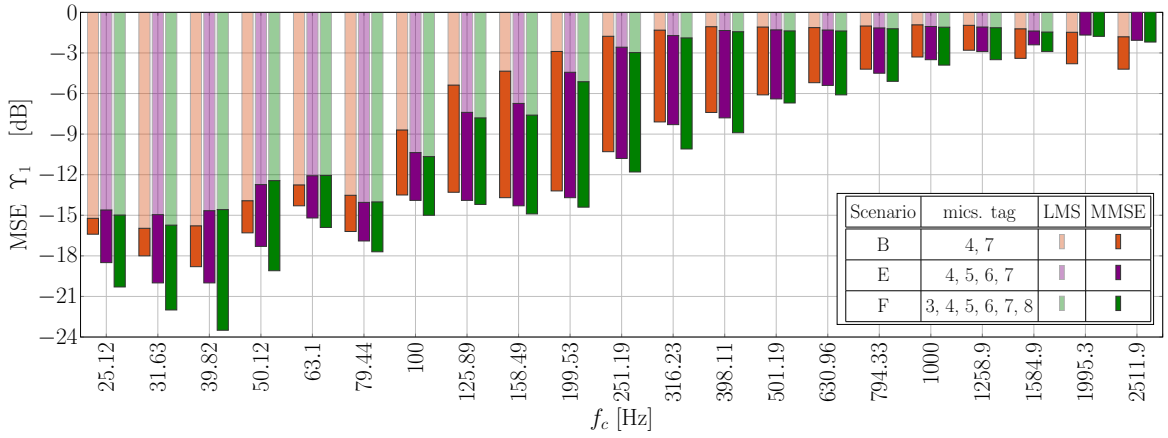


Figure H.18: Performance comparison in terms of normalized MSE as a function of 1/3-octave bands between adaptive (LMS) and fixed (MMSE) observation filter estimation approaches at the passenger's left virtual position for the best scenarios, selected among those with the same number of monitoring microphones, in the `Headrest` setup.

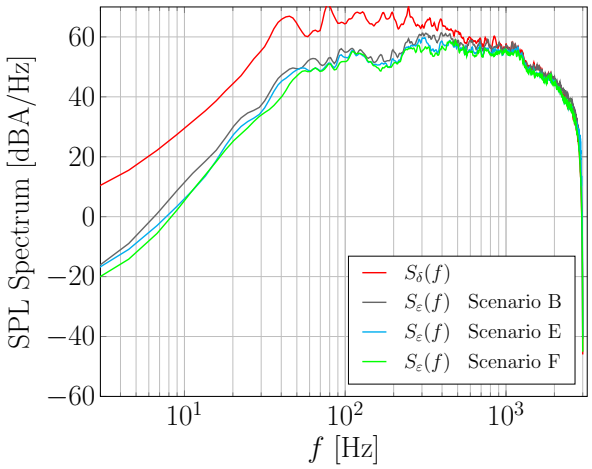


Figure H.19: SPL spectra of $\delta[n]$ and $\varepsilon[n]$ comparison for the best scenarios, selected among those with the same number of monitoring microphones, for the passenger's left microphone virtualization by means of MMSE algorithm with subcompact car experimental measurements and `Headrest` setup.

Conclusions on VMT

In Part II, a deep investigation on VMT for ANC systems in automotive applications is presented.

The goal of ANC is to generate a quiet zone around a confined area affected by a disturbing acoustic signal detected by a monitoring microphone. Since most of the ANC commercial solutions create a quiet zone around the error microphone, having the capacity to weaken the disturbance signal at a different position may be very useful. Especially for the automotive industry, in which the error microphone positioning is usually constrained by the car producer, physically reaching the desired quiet zones may be very complicated. Thus, it is possible to distinguish between the physical error microphones that are aimed to track disturbing noise, known as monitoring microphones, and the position in which the noise cancellation is targeted, e.g., near the driver/passenger's ears, called "virtual microphone. In order to effectively implement the ANC system with VMT, the estimation of the acoustic channels between monitoring and virtual microphones, referred to as observation filter, is necessary. In particular, a preliminary identification stage in which a physical microphone is momentarily installed at the desired virtual position is required in order to acquire the disturbing signal for observation filter estimation.

In Chapter 7, the estimate of the observation filters is assumed ideal. Hence, the mathematical backgrounds on both adaptive and fixed ANC systems with VMT are given. Chapter 8 deals with the algorithms used for the observation filter estimation and relevant issues. More precisely, In Section 8.1, the causality issue is presented. The mathematical derivation of the observation filter estimation, by adopting an adaptive approach, by means of LMS and RLS algorithms, or a fixed one through the MMSE, is given in Section 8.2. Finally in Section 8.3, the considered test on the observation filters robustness is presented.

Several and different signals for microphone virtualization are presented and analyzed in Chapter 9:

- Indirect measurements (realistic virtual microphone signals are unavailable) in Section 9.2
  - Realistic car signals filtered through synthetic filters in Section 9.2.1
  - White gaussian noise filtered through experimental filters in Section 9.2.2
- Direct measurements (realistic virtual microphone signals available) in Section 9.3
  - Sedan experimental setup in Section 9.3.1
  - Luxury car experimental setup in Section 9.3.2
  - Subcompact car experimental setup in in Section 9.3.3.

For all the considered cases, simulations were performed using MATLAB and Simulink. In particular, the indirect measurements show that both adaptive and fixed observation filter estimation approaches perform well, since they guarantee an efficient estimation of the observation filter. An empirical search of the optimal value of delay in order to obtain a causal system is performed when experimental filters are employed. Results show that the performance of the observation filter estimated by LMS and RLS algorithms is almost equal. Moreover, a significant degradation of the system performance can be observed, due to the increased complexity of the experimental filters with respect to the synthetic ones.

Regarding the direct measurements, disturbance signals are acquired from experimental measurement campaigns performed in part by ASK Industries on realistic car interiors, namely a sedan, a luxury and a subcompact car. The observation filter estimation accuracy is assessed in terms of SPL and MSE. Experimental results show how the strategic microphones placement plays a key role to obtain a sufficient spectral coherence of the monitoring and virtual microphones. Within the low frequency regime, the estimation approaches, i.e., LMS/RLS and MMSE, perform well. In general, it is shown that the MMSE algorithm performs well since good wide-band performance is obtained. Furthermore, the MMSE algorithm exhibits a remarkable robustness in case of driving scenario mismatch. Regarding the subcompact car experimental measurements, a pragmatic and heuristic method to determine the positioning and the

minimum number of monitoring microphones, required within the car cabin in order to effectively virtualize a specific position, is suggested.

Finally, in Chapter 10, a TCN for microphone virtualization is presented. In particular the use of DNN to perform the virtualization task is investigated. The obtained numerical results are most likely restrained by the limited size of the data-sets, yielding a large and robust microphone virtualization only in the low-frequency range. Furthermore, these datasets were not recorded in sufficiently extensive driving scenarios, therefore they lack in diversity. In fact, notoriously, DNNs require a considerably large amount of diversified data to be properly trained in order to generalize a complex task. Moreover, larger and deeper networks may be possibly required to obtain better results, also given the complexity of the task at hand. The employment of larger networks requires increased computational time and hardware requirements, but also the already mentioned need of a very large data-set. These observations suggest the requirement of a larger data-set to be used in the context of microphone virtualization, which could be addressed as a future work on the subject.

Bibliography

[1] P. N. Samarasinghe, W. Zhang and T. D. Abhayapala, "Recent advances in active noise control inside automobile cabins: Toward quieter cars," *IEEE Sign. Proc. Mag.*, vol. 33, no. 6, pp. 61–73, 2016.

[2] P. Peretti, S. Cecchi, L. Romoli and F. Piazza, "Adaptive Feedback Active Noise Control for Yacht Environments," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 2, pp. 737–744, 2014.

[3] S. M. Kuo and D. Morgan, "Active noise control: a tutorial review," *Proceedings of the IEEE*, vol. 87, no. 6, pp. 943–973, 1999.

[4] ——, *Active Noise Control Systems: Algorithms and DSP Implementations*. New York: John Wiley & Sons, 1995.

[5] A. Costalunga, C. Tripodi, L. Ebri, M. Vizzaccaro, L. Cattani, E. Ugolotti and T. Nili, "Experimental results on active road noise cancellation in car interior," in *Audio Engineering Society Convention 144*, Milan, May 2018.

[6] M. Pawelczyk, "Adaptive Noise Control Algorithms for Active Headrest System," *Control Engineering Practice*, vol. 12, pp. 1101–1112, 2004.

[7] P. Lueg, "Process of Silecing Sound Oscillations," United States Patent no. 2043416A, Jun. 1936.

[8] L. J. Fogel, "Apparatus for Improving Intelligence Under High Ambient Noise Levels," United States Patent no. 2966549A, Dec. 1960.

[9] W. Meeker, W. A. D. Center, and R. C. of America, *Active Ear Defender Systems: Development of a Laboratory Model*, ser. WADC

technical report. Wright Air Development Division, Air Research and Development Command, United States Air Force, 1959. [Online]. Available: https://books.google.it/books?id=SSSeXjSL7bIC

[10] Bose Corporation, "Bose History," https://boseaviation.com/about/history/, Online — accessed 13-Dec-2021".

[11] Sennheiser Corporation, "Sennheiser History," https://en-de.sennheiser.com/news-sennheisers-noisegard-active-noise-reduction, Online — accessed 13-Dec-2021".

[12] Y.-H. Lee and A. Nasiri, "Real time active noise control of engine booming in passenger vehicles," in *SAE Technical Paper Series*. SAE International, apr 2007. [Online]. Available: https://doi.org/10.4271%2F2007-01-0411

[13] Harman Cardon, "Halo Sonic, Engine Order Controlo," https://www.halosonic.co.uk/engine-order-cancellation/, Online — accessed 13-Dec-2021".

[14] S.-H. Oh, H.-s. Kim, and Y. Park, "Active control of road booming noise in automotive interiors," *The Journal of the Acoustical Society of America*, vol. 111, no. 1, pp. 180–188, 2002.

[15] D. Thompson and J. Dixon, "Chapter 6: Vehicle noise," in *Advanced Applications in Acoustics, Noise and Vibration*, F. Fahy and J. Walker, Eds. Spon Press, 2004, pp. 236–291.

[16] S. J. Elliott and P. A. Nelson, "Active Noise Control," *IEEE Signal Processing Magazine*, vol. 10, no. 4, pp. 12–35, 1993.

[17] B. F. Boroujeny, *Adaptive Filters: Theory and Applications*, 2th ed. New York, NY, USA: John Wiley & Sons, 1998.

[18] G. Zhaoa, N. Alujević, B. Depraetere, G. Pinte, J. Swevers and P. Sas, "Experimental study on active structural acoustic control of rotating machinery using rotating piezo-based inertial actuators," *Journal of Sound and Vibration*, vol. 348, pp. 15–30, 2015.

[19] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, 1 ed. New York, NY, USA: John Wiley & Sons, 1996.

[20] D. A. Bies and C. H. Hansen, *Engineering Noise Control: Theory and Practice*. CRC Press, 1996.

[21] A. Thompson and B. N. Taylor, *Guide for the Use of the International System of Units (SI)*. National Institute of Standard and Technology, Special Pubblication 811, 2008.

[22] A. N. Rimell, N. J. Mansfield and G. S. Paddan, "Design of digital filters for frequency weightings (A and C) required for risk assessments of workers exposed to noise," *Industrial Health*, vol. 53, no. 1, pp. 21–27, Sep. 2014.

[23] R. L. Rabiner and R. W. Schafer, "Introduction to Digital Speech Processing," in *Foundations and Trends in Signal Processing*, R. M. Gray, Ed. Hanover, MA, USA: Now Publishers, 2007, vol. 1, no. 1–2, pp. 1–194.

[24] S. Kay, *Modern Spectral Estimation: Theory and Application*. Prentice Hall, Mar. 1999.

[25] T. J. Sutton, S. J. Elliott, A. M. McDonald and T. J. Saunders, "Active control of road noise inside vehicles," *Noise Control Engineering J.*, no. 42, pp. 137–147, 1994.

[26] N. Zafeiropoulos, A. T. Moorhouse, M. Ballatore and A. Mackay, "Active Control of Road Noise: The Relation Between Reference Sensor Locations and the Effect on the Controller's Performance," University of Salford, Acoustics Research Centre, Tech. Rep., Jul. 2015.

[27] S. J. Elliott and T. J. Sutton, "Performance of Feedforward and Feedback Systems for Active Control," *IEEE Trans. on Speech and Audio Proc.*, vol. 4, no. 3, pp. 214–223, May 1996.

[28] W. Gan, S. Mitra, and S. Kuo, "Adaptive feedback active noise control headset: implementation, evaluation and its extensions," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 3, pp. 975–982, 2005.

[29] M. Rupp and A. H. Sayed, "Modified FxLMS Algorithms with Improved Convergence Performance," *Conference Record - Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1255–1259, 1995.

[30] G, Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms and approximate inverses," *IEEE Trans. Automatic Control*, vol. AC-26, no. 4, pp. 301–320, 1981.

[31] S. Liebich, C. Anemüller, P. Vary, P. Jax, D. Rüschen S. Leonhardt, "Active noise cancellation in headphones by digital robust feedback control," *European Signal Processing Conference*, pp. 1843–1847, 2016.

[32] M. Bai and D. Lee, "Implementation of an active headset by using the H_∞ robust control theory," *Journal of Acoustic Society of America*, vol. 104, no. 4, pp. 2184–2190, October 1997.

[33] H. Sano, T. Inoue, A. Takahashi, K. Terai and Y. Nakamura, "Active control system for low-frequency road noise combined with an audio system," *IEEE Trans. Speech and Audio Proc.*, vol. 9, no. 7, October 2001.

[34] B. Rafaley and S. J. Elliott, "H_2/H_∞ Active control of sound in a headrest: design and implementation," *IEEE Trans. Control Tech.*, vol. 7, no. 1, January 1999.

[35] L. Zhang L. Wu and Q. Xiaojun, "An intuitive approach for feedback active noise controller design," *Applied Acoustics*, vol. 74, no. 1, pp. 160–168, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.apacoust.2012.07.006

[36] J. C. Doyle, B. A Francis and A. R. Tannenbaum, *Feedback Control Theory*. Prentice Hall Professional Technical Reference, 1991.

[37] G. Stein, "Respect the unstable," *IEEE Control Systems Magazine*, vol. 23, no. 4, pp. 12–25, Aug 2003.

[38] A. Siswanto, C.Y. Chang and S. M. Kuo, "Active Noise Control for Headrests," *2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2015*, vol. 1, no. December, pp. 688–692, 2016.

[39] C. D. Peterson, R. Fraanje, B. S. Cazzolato, A.C. Zander and C. H. Hansen, "A Kalman filter approach to virtual sensing for active noise control," *Mech. Syst. Signal Process.*, vol. 22, no. 2, pp. 490–508, 2008.

[40] J. Garcia-Bonito, S. J. Elliott and C. C. Boucher, "Generation of zones of quiet using a virtual microphone arrangement," *J. Acoust. Soc. Am.*, vol. 101, no. 6, pp. 3498–3516, 1997.

[41] A. Roure, A. Albarrazin and S. Douglas, "The remote microphone technique for active noise control," in *Proceedings of Active Control of Sound and Vibration*, vol. 2. Institute of Noise Control Engineering of the USA, 1999, pp. 1233–1244.

[42] D. J. Moreau, J. Ghan, B. S. Cazzolato, and A. C. Zander, "Active noise control in a pure tone diffuse sound field using virtual sensing," *The Journal of the Acoustical Society of America*, vol. 125, no. 6, pp. 3742–3755, 2009.

[43] D. Das, D. Moreau, and B. Cazzolato, "Performance evaluation of an active headrest using the remote microphone technique," *Australian Acoustical SocietyConference 2011, Acoustics 2011: Breaking New Ground*, pp. 1–7, November 2011.

[44] S. Elliott, C. K. Lai, T. Vergez, and J. Cheer, "Robust stability and performance of local active control systems using virtual sensing," *Proceedings of the 23rd. International Congress on Acoustics*, pp. 61–68, 2019.

[45] Lichuan Liu, S. M. Kuo, and M. Zhou, "Virtual sensing techniques and their applications," in *2009 International Conference on Networking, Sensing and Control*, Okayama, Japan, May 2009, pp. 31–36.

[46] S. Elliott, "Virtul microphone arrangement for local active sound control." in *Paper presented at 1st Int. Conf. Motion and Vibration Control (Session: Control of Sound 1)*, 1992, pp. 1027–1031.

[47] B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, April 1988.

[48] P. S. S. Ahamed and P. Duraiswamy, "Virtual sensing active noise control system with 2d microphone array for automotive applications," in *2019 6th International*

Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, March 2019.

[49] T. Xiao, X. Qiu, and B. Halkon, "Ultra-broadband local active noise control with remote acoustic sensing," *Springer Nature Scientific Reports*, vol. 10, no. 1, November 2020. [Online]. Available: http://dx.doi.org/10.1038/s41598-020-77614-w

[50] 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.

[51] D. Shi, B. Lam, and W.-s. Gan, "Analysis of multichannel virtual sensing active noise control to overcome spatial correlation and causality constraints," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 8499–8503.

[52] J. Zhang, S. J. Elliott, and J. Cheer, "Robust performance of virtual sensing methods for active noise control," *Mechanical Systems and Signal Processing*, vol. 152, p. 107453, 2021.

[53] W. Jung, S. J. Elliott, and J. Cheer, "Local active control of road noise inside a vehicle," *Mechanical Systems and Signal Processing*, vol. 121, pp. 144–157, 2019.

[54] F. M. Fazi and P. Nelson, "The ill-conditioning problem in sound field reconstruction," *Journal of the Audio Engineering Society*, October 2007.

[55] A. Opinto, M. Martalò and R. Raheli, "Control of the Acoustic Field in a Vehicle Cabin," University of Parma, Department of Engineering and Architecture, Tech. Rep. 2, Dec 2020.

[56] Zoom Corporation, "Zoom F8," https://zoomcorp.com/en/jp/handheld-video-recorders/field-recorders/f8/, Online — accessed 20-Sept-2021".

[57] A. Opinto, M. Martalò, A. Costalunga, N. Strozzi, C. Tripodi and R. Raheli, "Experimental Results on Observation Filter Estimation for Microphone Virtualization," in *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, September 2021, pp. 1–7.

[58] H. Purwins, B. Li, T. Virtanen, J. Schlüter, S. Chang, and T. Sainath, "Deep learning for audio signal processing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 2, pp. 206–219, April 2019.

[59] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Springer Nature Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, December 2020.

[60] C. C. Aggarwal, *Neural Networks and Deep Learning: A Textbook*, 1st ed. Springer Nature, 2016.

[61] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks for action segmentation and detection," November 2016. [Online]. Available: https://arxiv.org/abs/1611.05267

[62] R. Straccia, "Temporal Convolutional Networks for Microphone Virtualization," Master's thesis, University of Parma – Department of Engineering and Architecture, Mar. 2021.

[63] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *27th International Conference on International Conference on Machine Learning*. Madison, WI, USA: Omnipress, June 2010, p. 807–814.

[64] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[65] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, pp. 193–202, 1980.

[66] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012. [Online]. Available: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

[67] "deepDreamImage function." [Online]. Available: https://www.mathworks.com/help/deeplearning/ug/visualize-features-of-a-convolutional-neural-network.html

[68] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks for action segmentation and detection," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1003–1012.

[69] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions," *World Scientific International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems (IJUFKS)*, vol. 6, pp. 107–116, April 1998.

[70] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Highway networks," November 2015. [Online]. Available: https://arxiv.org/abs/1505.00387

[71] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," December 2015. [Online]. Available: http://arxiv.org/abs/1512.03385

Acknowledgments

I would like to express my gratitude to Prof. Nicolò Cavina, our PhD coordinator, who made this interesting PhD program possible. I wish to show my honest appreciation to Prof. Stefano Selleri, co-advisor of my thesis, who allowed me to take part in this academic opportunity. Thanks for his suggestions. Special thanks to my advisor, Prof. Angelo Farina, the human dispenser on acoustics, for his guidance on the project, but also for his fundamental suggestions on professional and job growth.

At the end of this incredible adventure in Parma which was the bachelor's, master's degree and finally the PhD, I would like to express my gratitude to several persons. Thanking persons, such as talking about yourself, is never an easy task because you may risk to be too little generous or too much. Moreover, this is not a simple acknowledgment to persons who I knew during this long university period, these persons which collaborated and worked with me they have not been (let's say) simple "colleagues" and I did not studied/worked/found solutions, for and with (let's say) simple "university professors" but all these guys represent the part of my family that live in Department of Engineering and Architecture, Information Engineering Unit, Scientific Offices - Building 2, University Campus, Parco Area delle Scienze, 181/A, I-43124 – Parma (PR) Italy. Hence, this is an hard task, and the only way to do this fairly is as honestly, as I am doing.

The first person I would like to say thank you is Davide. In this case, the words are really, enormously difficult to choose even for a poet or a novelist and a series of many words would not be enough... For this reason, I limit to say thank you Davide, thank you for being one of the best person that I ever knew. Goodbye my friend, wherever you are.

Of course, thanks to my professor, my mentor and my advisor since 2014, Prof. Riccardo Raheli who supported me and bear my personality that sometimes was too impulsive and fiery (volcanic by nature, 100% made in Catania). He taught me not only the mathematical and theoretical skills but the dedication to the work. Even when the conditions were not so easy (especially during my first year of PhD but also during the pandemic), he continued to work and help me. For this reason, he is a source of inspiration.

If the stated "dietro ogni grande uomo c'è una grande donna" that, in english, can be translated in "behind a great man there is a great woman" holds, "behind a great professor there is a great researcher" holds too. This work would not have been possible without the help and support of Prof. Martalò. In fact, his contribution was fundamental, having guided and encouraged me to develop the project by giving me always significant moral support. Thank you Marco.

Thanks to Prof. Ferrari for helping me and following me also in other projects.

Thanks to the other guys of Pal. 2, my favorite girl-office room mate (not because she is the only one) namely, Veronica Mattioli. I enjoyed to stay with her in the office watching video regarding food, cakes, kittens and other "pucciosi" animals. Well... we also spent a bit of time working. We spent a lot of time on kittens, doggies and Plutino Caciotta. The rest we just squandered.

Thank you Luca, Laura, Gabriele and Gaia. I will not forget the sushi lunch breaks spent with you and the quantity of food we are able to gobble. Obviously, thanks to Franceso Denaro, my personal counselor regarding TFR, partita IVA, RAL and money stuff. Cognomen omen. Thanks to the guys of Pal. 4 and Pal. 6, Danilo, Federico, Daniel, Costante, Anna, Marco and Thomas.
Regarding the other guys outside the University who love me, thanks very much for stay always by my side. Thanks to (randomly listed) Lu Diegu, Pollo Gau, Elo Elo, Claudia "the mommy", Miriam "the recovered", Shennon "the limit", my cousin Valeria (fai cagare), Peppe, Franca, Suppa, Tommane and Francisca. From Catania with the van, thanks the most important persons of my life. Thanks to Bam Bam Ciccio "the

triple 20", Sappa "the game master", Gullo "la Esse", Jamba "never like before", Peppe Fuerte "de la mala suerte dammi un Peppe Fuerte" and Mimmuzzo. All of you guys, are very important to me and you represent my family.
Thank you Maximum Likelihood (you know what I mean).

Thank you city of Parma. Goodbye.

Of course, special thanks to my mum, daddy, Rossi and Giusy ❤.